



**The *in silico* Investigation of Pharmacological Targets of the Zika
Virus:
Insights into the Structural Characteristics of the NS5 and NS3
Proteins from Atomistic Molecular Simulations.**

Miss Pritika Ramharack

2017

A thesis submitted to the School of Health Sciences, University of KwaZulu
Natal, Westville, in fulfilment for the degree of Doctor of Philosophy

**The *in silico* Investigation of Pharmacological Targets of the Zika Virus:
Insights into the Structural Characteristics of the NS5 and NS3 Proteins
from Atomistic Molecular Simulations.**

Miss P Ramharack

211510436

2016

A thesis submitted to the School of Pharmacy and Pharmacology, Faculty of Health Sciences, University of KwaZulu-Natal, Westville, for the degree of Doctor of Philosophy.

This is the thesis in which the chapters are written as a set of discrete research publications, with an overall introduction and final summary. Typically these chapters will have been published in internationally recognized, peer-reviewed journals.

This is to certify that the contents of this thesis are the original research work of Miss P Ramharack.

As the candidate's supervisor, I have approved this thesis for submission.

Supervisor:

Signed: *Mahmoud E. Soliman*

Name: Prof Mahmoud Soliman

Date: 31 May 2017

PREFACE

This thesis is divided into **eight chapters**, including this one:

Chapter 1:

This is an introductory chapter that addresses the background, rationale and relevance of the study as well as the proposed aim and objectives. The general outline and structure of the thesis concludes this chapter.

Chapter 2:

This chapter provides a comprehensive literature review on the ZIKV epidemic and the urgent research currently underway toward the development of FDA approved inhibitors of the virus. Included in this chapter is the epidemiology, historical background, life cycle, viral diagnostics, modes of transmission, ZIKV-linked neurological diseases, viral characteristics (mechanistic and structural), viral/host drug targets, specifically the NS5 and NS3 protein and the design of potential inhibitors in ZIKV rational drug design and discovery.

Chapter 3:

This chapter conceptualizes computer-aided drug design by discussing a various molecular modeling and molecular dynamic techniques and applications. The computational tools needed to investigate comparative enzymatic structural/conformational characteristics as well as methods used to analyze binding affinity are elucidated upon.

Chapter 4: (Published work- this chapter is presented in the required format of the journal and is the final version of the accepted manuscript)

This chapter demonstrates a unique route map entitled “*Zika virus drug targets: a missing link in drug design and discovery – a route map to fill the gap*”, demonstrating potential drug targets, strategies for design and computational software available to design a homology model. Also presented is a 3D homology model of the ideal ZIKV target, the non-structural protein 5 in which the active binding sites of each domain of the protein were identified and structure-based virtual screening allowed for the identification of possible NS5 RdRp small molecule inhibitors. This article has been published in *RSC Advances* (IF = 3.289). It should be noted that the publication was completed prior to the release of the 3D crystal structure of the NS5 methyltransferase and RNA-dependent RNA-polymerase.

Chapter 5: (Published work- this chapter is presented in the required format of the journal and is the final version of the submitted manuscript)

This chapter investigates the second objective of the thesis and is entitled “*Zika Virus NS5 Protein Potential Inhibitors: An Enhanced In silico Approach in Drug Discovery.*” The study implements an optimized and proven screening technique in the discovery of two potential small molecule inhibitors of ZIKV MTase and RdRp. This *in silico* “per-residue energy decomposition pharmacophore” virtual screening approach will be critical in aiding scientists in the discovery of effective inhibitors of ZIKV targets. This article has been published in *Journal of Biomolecular Structure and Dynamics* (IF = 2.3).

Chapter 6: (Published work- this chapter is presented in the required format of the journal and is the final version of the submitted manuscript)

This chapter, “*Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery*”, assesses the third objective of the thesis: to identify the structural properties of the ZIKV NS3 Helicase when bound to ATP-competitive inhibitor, NITD008. In this study, comparative molecular dynamic simulations were employed for Apo and bound protein to demonstrate the molecular mechanism of the Helicase, thus assisting in the design of effective inhibitors against this detrimental viral target. The article has been published in *RSC Advances* (IF = 3.289).

Chapter 7: (Published work- this chapter is presented in the required format of the journal and is the final version of the submitted manuscript)

This chapter is entitled “*Characterizing the Conformational Features and Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors*”, elucidates on binding landscape of the ATPase and ssRNA site by demonstrating the chemical characteristics of potent *flavivirus* lead compounds, Lapachol, HMC-HO1 α and Ivermectin at the respective NS3 Helicase binding site. This article has been published in *Future Virology* (IF = 0.886).

Chapter 8:

This is the final chapter that proposes future work and concluding remarks.

ABSTRACT

The re-emerging Zika virus has evolved into a catastrophic epidemic during the past year, with an estimated 1.5 million reported cases of Zika infections worldwide, since the 2015 outbreak in Brazil. The virus has received considerable attention during 2016 with a flood of new discoveries, from evolving modes of viral transmission to viral-linked neurological disorders, unique specificity to host cells and increasing mutation rates. However, prior to the devastating 2015 outbreak in Brazil, the virus was classified as a neglected pathogen similar to Dengue and the West Nile virus.

Despite the wide-scale research initiative, there is still no cure for the virus. There are currently vaccine clinical trials that are on-going but there has not been a breakthrough with regard to small molecule inhibitors. A lot of experimental resources have been allocated to repurposing FDA-approved drugs as possible inhibitors, however, even some of the most potent *flavivirus* inhibitors have adverse toxic effects. The first crystal structure of the Zika virus was released in May 2016 and since then, six viral protein structures have been made available. Due to this lack in structural information, there is little known regarding the structural dynamics, active binding sites and the mechanism of inhibition of ZIKV enzymes.

This study delves into the structural characteristics of three of the most crucial enzymatic targets of the Zika virus, the NS5 RNA-dependent RNA polymerase and Methyltransferase as well as the NS3 Helicase. With emerging diseases, such as ZIKV, computational techniques including molecular modeling and docking, virtual screening and molecular dynamic simulations have allowed chemists to screen millions of compounds and thus funnel out possible lead drugs. These *in silico* approaches have warranted Computer-Aided Drug Design as a cost-effective strategy to fast track the drug discovery process.

The above techniques, amongst numerous other computational tools were employed in this study to provide insights into conformational changes that elucidate potential inhibitory mechanisms, active site identification and characterization and pharmacophoric features leading to promising small molecule inhibitor candidates.

The first study (Chapter 4), provided a comprehensive review on potential host/viral targets as well as provided a concise route map depicting the steps taken toward identifying potential inhibitors of drug

targets when no crystal structure is available. A homology model case study, of the NS5 viral protein, was also demonstrated.

The second study (Chapter 5) used the validated NS5 homology model to investigate the active sites at both the RNA-dependent RNA polymerase and Methyltransferase domains and subsequently employ a generated pharmacophore model to screen for potential inhibitors.

Chapter 6 reports the third study, which investigates the structural dynamics and in turn, the possible mechanism of inhibition of the ZIKV NS3 Helicase enzyme when bound to ATP-competitive inhibitor, NITD008. The study also provides insight on the binding mode at the ATPase active site, thus assisting in the design of effective inhibitors against this detrimental viral target.

Chapter 7 maps out the binding landscape of the ATPase and ssRNA site by demonstrating the chemical characteristics of potent *flavivirus* lead compounds, Lapachol, HMC-HO1 α and Ivermectin at the respective NS3 Helicase binding sites.

This study offers a comprehensive *in silico* perspective to fill the gap in drug design research against the Zika virus, thus giving insights toward the structural characteristics of pivotal targets and describing promising drug candidates. To this end, the work presented in this study is considered to be a fundamental platform in the advancements of research toward targeted drug design/delivery against ZIKV.

DECLARATION I -PLAGIARISM

I, Pritika Ramharack, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a. Their words have been re-written but the general information attributed to them has been referenced.
 - b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

A detailed contribution to publications that form part and/or include research presented in this thesis is stated (include publications submitted, accepted, in press and published).

Signed: **P.Ramharack**

DECLARATION II- LIST OF PUBLICATIONS

1. Pritika Ramharack and Mahmoud E.S. Soliman (2016) Zika Virus Drug Targets: A Missing Link In Drug Design And Discovery – A Route Map To Fill The Gap, *RSC Advances*, 6, 68719-68731. (*Published*)

Contribution:

Pritika Ramharack: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Mahmoud E.S. Soliman: Supervisor

Appendix A: Pdf version of the publication

2. Pritika Ramharack and Mahmoud E.S. Soliman (2017) Zika Virus NS5 Protein Potential Inhibitors: An Enhanced In silico Approach in Drug Discovery, *Journal of Biomolecular Structure & Dynamics*. (*Published*)

Contribution:

Pritika Ramharack: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Mahmoud E.S. Soliman: Supervisor

Appendix B: Pdf version of the publication

3. Pritika Ramharack, Sofiat Oguntade, Mahmoud E.S. Soliman (2017) Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery, *RSC Advances*, 7, 22133-22144. (*Published*)

Contribution:

Pritika Ramharack: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Sofiat Oguntade: contributed to the project by assisting in data collation

Mahmoud E.S. Soliman: Supervisor

Appendix C: Pdf version of the publication

4. Sofiat Oguntade, Pritika Ramharack and Mahmoud E.S. Soliman (2017) Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors, *Future Virology*. (In Press)

Contribution:

Pritika Ramharack: contributed to the project by performing manuscript preparation and writing.

Sofiat Oguntade: contributed to the project by performing experimental work

Mahmoud E.S. Soliman: Supervisor

RESEARCH OUTPUT

A- LIST OF PUBLICATIONS

1. Pritika Ramharack and Mahmoud E.S. Soliman (2016) Zika Virus Drug Targets: A Missing Link In Drug Design And Discovery – A Route Map To Fill The Gap, *RSC Advances*, 6, 68719-68731.
2. Pritika Ramharack and Mahmoud E.S. Soliman (2017) Zika Virus NS5 Protein Potential Inhibitors: An Enhanced *in silico* Approach in Drug Discovery, *Journal of Biomolecular Structure & Dynamics*, 2017, 1-16.
3. Pritika Ramharack, Sofiat Oguntade, Mahmoud E.S. Soliman (2017) Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery, *RSC Advances*, 7, 22133-22144.

- IN PRESS

1. Sofiat Oguntade, Pritika Ramharack and Mahmoud E.S. Soliman (2017) Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors, *Future Virology*.

B- CONFERENCES

1. Poster Presentation “Zika Virus NS5 Protein Potential Inhibitors: An Enhanced *in silico* Approach in Drug Discovery”- International Conference on Medicinal and Pharmaceutical Chemistry, Dubai, UAE, 5-7 December 2016.

ACKNOWLEDGEMENTS

My family:

To my parents, my love and respect for you both are as great as the universe. Your trust and prayers have seen me through sleepless nights and painful days away from you. You provided for me even if it meant you having nothing for yourself. Your sacrifices will never be forgotten.

Professor M.E.S Soliman:

I would like to thank you for accepting me into your prestigious department and for your unwavering support and trust throughout this degree.

The Molecular Modeling and Drug Design Group:

Thank you for all for the laughs, debates and for assistance throughout the year. The memories shall be forever a reminder of the friendships made.

My Loved ones and Friends:

My friends who have become my family through the years. Every laugh, cry and smile we have shared is treasured. I could not have asked for better group of people to share these years with. Nishall, your daily support and well wishes are what kept me motivated through even the rainiest days, thank you.

University of KwaZulu-Natal College of Health Sciences and National Research Foundation

Financial support as postgraduate student.

LIST OF ABBREVIATIONS

Å	Amperes
α	Alpha
aMD	Accelerated molecular dynamics
ATP	Adenosine Triphosphate
β	Beta
CADD	Computer-Aided Drug Design
CNS	Central nervous system
CST	Castanospermine
DCCM	Dynamic Cross Correlation
DNA	Deoxyribonucleic acid
DNJ	deoxynojirimycin
FDA	Food and Drug Administration
Flu	Influenza
ΔG	Free Binding energy
GAFF	General amber Force Field
GTP	Guanine Triphosphate
IFN-1	Interferon-1
JEV	Japanese Encephalitis Virus
K	Kelvin
LBVS	Ligand-based virtual screening
MC	Monte Carlo
MD	Molecular Dynamics
MM	Molecular Mechanics

MM/GBSA	Molecular Mechanics/Generalized Born Surface Area
MM/PBSA	Molecular Mechanics/Poisson-Boltzmann Surface Area
MTase	Methyltransferase
NMR	Nuclear Magnetic Resonance
ns	nanoseconds
NS	Non-structural
ORF	Open Reading Frame
PBVS	Pharmacophore-based virtual screening
PCA	Principal Component Analysis
PDB	Protein data bank
PME	Particle-mesh Ewald method
PRED	Per Residue Decomposition
prM	Pre-Membrane
ps	Picoseconds
QM	Quantum Mechanics
RdRp	RNA Dependent RNA polymerase
RESP	Restrained Electrostatic Potential
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RNA	Ribonucleic Acid
RoG	Radius of Gyration
RSCB	Research Collaboratory for Structural Bioinformatics
RT-PCR	Real Time Polymerase Chain Reaction

SAM	<i>S</i> -adenosylmethionine
SASA	Solvent accessible surface area
SBVS	Structure-based virtual screening
siRNA	Silencing ribonucleic acid
SPCS	Signal peptidase complex
ssRNA	Single-stranded ribonucleic acid
STAT2	Signal Transducer And Activator Of Transcription 2
TLR3	Toll-like receptor3
UTR	Un-Translated Region
VS	Virtual screening
WHO	World Health Organization
ZIKV	Zika Virus
3D	Three-Dimensional

LIST OF AMINO ACIDS

Three Letter Code	Amino Acid
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
Cys	Cysteine
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

LIST OF FIGURES

Figure 2.1: Time-line demonstrating the historical outbreaks of ZIKV and the increase in virulence with each strain (Image prepared by author).	12
Figure 2.2: Zika Virus Life cycle from vector to host transmission and replication (Adapted from (Screaton <i>et al.</i> 2015)).	14
Figure 2.3: Cleaved ZIKV polyprotein demonstrating available protein crystal structures (PDB codes: 5IY3, 5JHM, 5JMT, 5T1V, 5KQR, 5UO4) (Prepared by Author).	15
Figure 2.4: Crystal structure of ZIKV NS5 protein. The protein comprises of three domains, the N-terminal Mtase domain (residues 1-262) (red), the C-terminal RdRp domain (residues 273-907) (green) and linker domain (residues 263-272) (blue), (PDB code: 5TFR) (Prepared by Author).	18
Figure 2.5: Crystal structure of ZIKV NS3 Helicase Protein depicting the three domains and two-active binding regions (blue) that form hydrophobic pockets for ATP and ssRNA binding (Prepared by Author).	20
Figure 3.1: The scientific domains in which Applications of Quantum and Molecular Mechanics fit into (Prepared by Author).	35
Figure 3.2: The Bohr Model demonstrated the atom to have a positively charged nucleus that was orbited by negatively charged electrons. (Prepared by Author).	37
Figure 3.3: graphical representation of a two-dimensional potential energy surface (PES) (University of California n.d.).	40
Figure 3.4: Diagramatic representation of the total potential energy function of a molecule, as mentioned above (Prepared by Author).	43
Figure 3.5: Diagramatic representation of thermodynamic cycle as justified in the equations above (Prepared by Author).	49
Figure 3.6: Summary of methods employed in this study (Prepared by Author).	54
Figure 4.1: Global reports of ZIKV transmission, infection and sporadic viral antibody reports prior to 2015, as of April 2016	63
Figure 4.2: Route map toward the <i>in silico</i> design of ZIKV inhibitors using the homology modeled viral NS5 protein.	71
Figure 4.3: Protocol for building a homology model in our laboratory.	72
Figure 4.4: Homology model of ZIKV NS5 protein.	74
Figure 4.5: The potential binding sites, identified by Site-hound (71), of ZIKV NS5 protein.	75
Figure 4.6: Docked Conformation of 2'-C-Methyladenosine with ZIKV NS5 RdRp (Binding affinity: -6.3kcal/mol).	79
Figure 5.1: A per residue energy decomposition-based approach outline applied	95

in the study.

Figure 5.2: Superimposition of homology model (yellow) with the newly released crystal structure (green- PDB code: 5TFR), showing their structural similarity and validating the model's use for subsequent analysis. 103

Figure 5.3: The steps taken toward creating the pharmacophore model from the MTase-BG323 complex. The yellow circles spotlight the pharmacophoric moieties that were chosen for the model, based on the highest contributing residues, depicted in the binding affinity graph. 104

Figure 5.4: RdRp-Ribavirin complex ligplot analysis- creating the pharmacophore model to virtually screen for new RdRp potential lead compounds. 105

Figure 5.5: The C α root mean square fluctuations (RMSF) of MTase-BG323 and MTase-ZINC64717952 during the molecular dynamic simulation. 108

Figure 5.6: MTase-ZINC64717952 complex interactions. 109

Figure 5.7: RdRp-ZINC39563464 complex interactions. 112

Figure 5.8: The stable C α root mean square fluctuations (RMSF) of RdRp-Ribavirin and RdRp-ZINC39563464 during the molecular dynamic simulation. 113

Figure 6.1: Cartoon and surface representation of the three domains of the ZIKV helicase and the two active-binding regions (yellow) that form profound hydrophobic cavities in the electrostatic surface area, allowing ATP and ssRNA to bind. 127

Figure 6.2: Energy contributions of the highest interacting residues at the ATPase active site. 137

Figure 6.3: Superimposed conformation of structurally similar NITD008 and ATP docked at ATPase site of ZIKV NS3 Helicase. 138

Figure 6.4: C- α backbone RMSD for NS3 Helicase Apo enzyme and NITD-complex conformation. The average C- α RMSD was calculated to be 3.62 Å and 3.77 Å, respectively. Increased fluctuations occurred at 47-52ns in the NITD008-complex. 140

Figure 6.5: The RMSF of Apo enzyme and NITD008-complex. The structural flexibility in domain I and II is highly attributed to the binding of NITD008 to the ATP-active site. 141

Figure 6.6: The radius of gyration (RoG) plot illustrating the difference in enzyme compactness of the NITD008-complex compared to the Apo enzyme. 142

Figure 6.7: Structural Flexibility of the P-Loop (196-203), RNA-binding loop (244-255), and the $_{310}$ Helix (339-348) along the trajectory. 144

Figure 6.8: Residue fluctuations at the P-Loop region. The Apo enzyme illustrates closing of the loop at the active site due to a vacant hydrophobic pocket. 146

Figure 6.9: Residue fluctuations at the “325-348” region. The Apo enzyme illustrates widening of the loops of the Apo enzyme.	148
Figure 6.10: Projection of Eigen values of the C α backbone, during 130 ns simulation, for Apo and NITD008-bound conformations of NS3 Helicase along the first two principal components.	150
Figure 7.1: Structure of NS3 ZIKV Helicase (PBD 5JMT), [21]. Domain 1 (blue: residue 175-332) and domain 2 (red: residue 333-481) are seen facing each other and domain 3 (green: residue 482-617) lying above the other 2 domains.	162
Figure 7.2: C-alpha RMSD backbone Plot for NS3 Helicase free and ligand bound conformations. The ligands Lapachol, HMC-HO1 α and Ivermectin are seen to stabilize the protein as compared to the fluctuating free protein.	171
Figure 7.3: (A) RMSF Plot for Lapachol, HMC-HO1 α and Ivermectin systems. Lapachol (0.88Å) showed a higher stability at the ATPase site compared to HMC-HO1 α (0.90Å) and Ivermectin (0.85Å) showed the most favorable stability of all the systems, (B) NS3 Helicase residue fluctuations at regions: 1- the “72-79” loop (Navy), 2- the “172-176” helix (Gold) and 3- the “409-411” loop (Magenta).	173
Figure 7.4: Radius of gyration Plot for Lapachol, HMC-HO1 α and Ivermectin systems when compared to the free protein.	175
Figure 7.5: Free energy decomposition and ligand-residue interaction network at the ATPase site of the Lapachol- NS3 Helicase system.	178
Figure 7.6: HMC-HO1 α docked into the ATPase site of Zika NS3 helicase, illustrating ligand-residue interactions and active-site residue energy contributions.	179
Figure 7.7: Free energy decomposition and ligand-residue interaction network at the ssRNA site of the Ivermectin- NS3 Helicase system.	180

LIST OF TABLES

Table 2.1: Overview of the currently PDB-deposited Crystal Structures of ZIKV NS5 and NS3 Proteins.	21
Table 2.2: Most Popular Repurposed Drugs as ZIKV Inhibitors	25
Table 4.1: Potential ZIKV target proteins	66
Table 4.2: Representation of top ten compounds docking to NS5 RdRp	77
Table 4.3: Physical Representation of top ten compounds displaying Molecular weight, xlogP, H-Bond Donors/Acceptors and Rotatable Bonds	79
Table 4.4: Proposed computational software used in ZIKV drug design	81
Table 5.1: Criteria summary of chosen templates used in Building the ZIKV NS5 homology model.	96
Table 5.2: Representation of the top three compounds bound to MTase and RdRp. The compounds, ZINC64717952 and ZINC39563464 showed the best binding affinity to MTase and RdRp respectively.	106
Table 5.3: The comparison of MTase's binding affinity with BG323 and ZINC64717952.	110
Table 5.4: The comparison of RdRp's binding affinity with Ribavirin and ZINC39563464.	111
Table 5.5: The comparison of drug likeness of the screened compounds compared to that of the experimental drugs against ZIKV.	114
Table 6.1: Summary of free binding Energy contributions to the NITD008-NS3 Helicase system.	136
Table 7.1: Calculated parameters for running accelerated molecular dynamics.	167
Table 7.2: 2D structure, docked complexes and validation of the docked complexes	170
Table 7.3: Binding free energy analysis (kcal/mol) for inhibitor- NS3 helicase complexes.	177

TABLE OF CONTENTS

PREFACE.....	III
ABSTRACT.....	V
DECLARATION I- PLAGIARISM.....	VII
DECLARATION II- LIST OF PUBLICATIONS.....	VIII
RESEARCH OUTPUT.....	X
ACKNOWLEDGEMENTS.....	XI
LIST OF ABBREVIATIONS.....	XII
LIST OF AMINO ACIDS.....	XV
LIST OF FIGURES.....	XVI
LIST OF TABLES.....	XIX
CHAPTER 1	
Introduction	
Background and Rational.....	1
Aim and Objectives.....	3
Novelty and Significance of Study.....	4
CHAPTER 2	
Background on the Zika Virus	
<i>Introduction.....</i>	<i>10</i>
<i>Epidemiology and Transmission.....</i>	<i>10</i>
<i>Characterizing ZIKV.....</i>	<i>12</i>
<i>Life cycle of ZIKV.....</i>	<i>13</i>
<i>Structural Characteristics of ZIKV.....</i>	<i>14</i>
<i>ZIKV Pathogenesis and Clinical Features.....</i>	<i>15</i>
<i>Rationale of ZIKV Enzymes as Potential Therapeutic Targets.....</i>	<i>16</i>
<i>NS5 Protein.....</i>	<i>17</i>
<i>NS3 Protein.....</i>	<i>19</i>
<i>Three-dimensional structures of ZIKV NS5 and NS3 proteins.....</i>	<i>20</i>
<i>Targeting Host Proteins in ZIKV Therapy.....</i>	<i>22</i>
<i>The Scientific Advancements of ZIKV anti-viral Therapy.....</i>	<i>24</i>
Preventative Antibodies and Vaccines.....	24
Small Molecule Inhibitors.....	25
CHAPTER 3	

Molecular Modeling and Computational Approaches to Biomolecular Structure and Drug Design

<i>Introduction.....</i>	34
<i>Principle of Quantum Mechanics.....</i>	36
<i>The Schrödinger Wave Function.....</i>	37
<i>The Born-Oppenheimer Approximation Theory.....</i>	38
<i>Potential Energy Surface as an Application of Quantum Mechanics.....</i>	39
The Principle of Molecular Mechanics.....	41
<i>Potential Energy Function.....</i>	41
The Principle of Molecular Dynamics.....	44
<i>Molecular Dynamics Post Analysis.....</i>	45
System Stability.....	46
<i>Thermodynamic Energy Calculations.....</i>	47
<i>Conformational Features of the System.....</i>	49
Other Computer-Aided Drug Design Techniques Utilized in the Study.....	51
<i>Homology Modeling.....</i>	51
<i>Molecular Docking.....</i>	52
<i>Virtual Screening.....</i>	53

CHAPTER 4

Zika Virus Drug Targets: A Missing Link In Drug Design And Discovery – A Route Map To Fill The Gap.....	60
Introduction.....	62
Overview of ZIKV Protein Assembly.....	64
Potential Biological Drug Targets Against ZIKV.....	65
<i>Viral Drug Targets.....</i>	65
<i>ZIKV Host Targets.....</i>	68
<i>in-silico studies conducted on ZIKV.....</i>	69
<i>in- silico route map toward the design and discovery of ZIKV inhibitors.....</i>	70
<i>A homology model for ZIKV NS5.....</i>	74
<i>Active site identification.....</i>	75
<i>Possible small molecule inhibitors of NS5 RdRp.....</i>	74
Proposed computational software that may be used in ZIKV drug design and discovery.....	76

Conclusion.....	82
CHAPTER 5	
Zika Virus NS5 Protein Potential Inhibitors: An Enhanced <i>In silico</i> Approach in Drug Discovery.....	89
<i>Introduction.....</i>	<i>91</i>
<i>Computational Methods.....</i>	<i>95</i>
<i>Results and Discussion.....</i>	<i>103</i>
<i>Conclusion.....</i>	<i>114</i>
CHAPTER 6	
Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery.....	123
<i>Introduction.....</i>	<i>125</i>
<i>Computational Methods.....</i>	<i>129</i>
<i>Results and Discussion.....</i>	<i>134</i>
<i>Conclusion.....</i>	<i>151</i>
CHAPTER 7	
Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors.....	158
<i>Introduction.....</i>	<i>160</i>
<i>Computational Methodology.....</i>	<i>165</i>
<i>Results and Discussion.....</i>	<i>169</i>
<i>Conclusion.....</i>	<i>180</i>
<i>Future Perspective.....</i>	<i>181</i>
CHAPTER 8	
Conclusion.....	189
Future Perspectives.....	190
APPENDIX.....	192

CHAPTER 1

1 Introduction

1.1 Background and Rational

The re-emerging Zika virus (ZIKV) has evolved into a catastrophic epidemic during the past year, with an estimated 1.5 million reported cases of ZIKV infections worldwide, since the 2015 outbreak in Brazil (Kollman *et al.* 2016). The arthropod-borne virus, together with Dengue, Japanese Encephalitis and West Nile virus, form part of the *flavivirus* genus, predominately found in the tropics. However, recent reports have evidenced new modes of transmission of ZIKV, including congenital, perinatal and sexual transmission, thus sanctioning the rapid spread of the virus on a global scale (Turmel *et al.* 2016; Tilak *et al.* 2016 Singh *et al.* 2016; Incicco *et al.* 2013; Gourinat *et al.* 2015; D'Ortenzio *et al.* 2016; Foy *et al.* 2011).

During ZIKV replication, the structural proteins (Capsid, membrane, pre-membrane and envelope protein) and non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) of the virus prove to be imperative for the replication of the RNA genome, virion assembly and invasion of the innate immune system (Noble *et al.* 2010; Mahfuz *et al.* 2014; Zanluca *et al.* 2016). By developing inhibitors against ZIKV-specific proteins, viral replication may be terminated with minimal adverse effects to the host. Of the ZIKV proteins, the NS3 and NS5 play a central role in viral RNA replication and maturation (Bollati *et al.* 2010). The NS3 protein is made up of two functional domains being the protease that is responsible for posttranslational cleavage of the nonstructural proteins at five sites on the protein chain and the helicase at the C-terminal being responsible for RNA binding and ATP hydrolysis (Chen *et al.* 2010; Murray *et al.* 2009; Kwong *et al.* 2005). The largest non-structural protein being the NS5 protein is made up of an N-terminal methyltransferase and a C-terminal RNA-dependent RNA polymerase that allow for 5'UTR capping and RNA synthesis, respectively (Murray *et al.* 2009; Bollati *et al.* 2010; Perera *et al.* 2008; Medin & Rothman 2016).

As of March 2016, International health associations announced ZIKV as a public health emergency based on growing evidence of the virus being linked to congenital neurological

diseases such as Guillain-Barre, cranial nerve dysfunction and Microcephaly (Broxmeyer & Kanjhan 2016; WHO 2016; Palomo 2016; Rasmussen *et al.* 2016). In response to the devastating consequences of antenatal infection, the scientific community invested significant research toward preventative and curative strategies, including vaccine and chemotherapeutic development. Although preventative clinical trials are under way, there are still no FDA approved small molecule inhibitors against the virus (Cohen 2016; Malone *et al.* 2016).

One of the most problematic tasks researchers have had to overcome is the ability of the virus to target neuronal cells, as inhibitors will not only need to be target-specific, effective and have minimal toxicity, but it will also have to pass through the blood-brain-barrier (Plourde & Bloch 2016; Anaya *et al.* 2016; Bayless *et al.* 2016; Olgarnier *et al.* 2016; Brault *et al.* 2016; Li *et al.* 2016; Nowakowski *et al.* 2016). Novel drug discovery and development, from design to the market, may take from anything between 10-20 years. With minimal literature available evidencing ZIKV's mechanism of action on host cells and evolving mutations of the virus, developing a novel drug that meets all the requirements of a ZIKV inhibitor may be laborious and costly.

Computer-Aided Drug Design is a cost-effective strategy to fast track the drug discovery process. Computational methods and resources may be implemented in most stages of drug discovery from identifying targets, to drug optimization and preclinical testing (Lu *et al.* 2012; Huang *et al.* 2010; Song *et al.* 2009; Anderson 2003). With emerging diseases, such as ZIKV, computational techniques including molecular modeling and docking, virtual screening, identification of pharmacophoric hot spots and molecular dynamic simulations allow chemists to screen millions of compounds to funnel out possible lead drugs which may then be validated experimentally. This strategy overcomes the concept of "shooting in the dark" with experimental screening, thus reducing the drug discovery time-line.

In this study, due to the lack of fundamental research in the previously neglected tropical disease, we have utilized key computational techniques to fill the gap in drug design research against ZIKV, thus giving insights toward viral drug targets and designing potential inhibitors against this new epidemic.

1.2 Aim and Objectives

The primary purpose of this thesis is to identify and characterize the principal target proteins of ZIKV and subsequently utilize Computer-Aided Drug Design techniques to investigate potential small molecule inhibitors against these proteins.

To accomplish this, the following objectives were outlined:

1. To create a concise route map depicting the steps taken toward identifying potential inhibitors of drug targets with no 3D crystal structure by:
 - 1.1. Providing a comprehensive review on ZIKV including potential viral/host targets.
 - 1.2. Creating a homology model and classifying the active sites of the essential ZIKV NS5 protein (prior to the release of the 3D crystal structure).
 - 1.3. Identification of potential inhibitors against the NS5 RNA-Dependent RNA polymerase from commercially chemical databases by performing structure-based virtual screening.
2. To utilize the *in silico* “Per-residue Energy Decomposition Pharmacophore” virtual screening technique to propose potential NS5 Methyltransferase (MTase) and RNA-Dependent RNA polymerase (RdRp) inhibitors. This may be achieved by:
 - 2.1. Performing molecular dynamic simulations to create molecular dynamic ensembles of potent *flavivirus* inhibitors in complex with Mtase and RdRp.
 - 2.2. Quantifying individual amino acid interactions towards total binding free energy based on the MM/GBSA approach, thus designing a pharmacophore model of each complex established from ligand-enzyme interactions.
 - 2.3. To subject the pharmacophore-based leads and search the Zinc Database for structure-based scaffolds against the respective enzyme to estimate their binding affinities.
 - 2.4. Validating the most favorable ligands by assessing the stability and binding free energy of each system following exposure to molecular dynamic simulations.
 - 2.5. Computing the physicochemical descriptors as well as predicting the pharmacokinetic properties and drug-like nature of the most favorable ligands.

3. To investigate the structural dynamics and in turn, the mechanism of inhibition of the ZIKV NS3 Helicase enzyme when bound to ATP-competitive inhibitor, NITD008. This will provide insight on the binding mode at the ATPase active site, thus assisting in the design of effective inhibitors against this detrimental viral target. These objectives are achieved by:
 - 3.1. Combining molecular docking with classical comparative molecular dynamic simulations of 100 ns for the free enzyme state as well as a NITD008-bound complex.
 - 3.2. Utilizing a wide variety of post-molecular dynamic analysis techniques to characterize the binding landscape of the enzyme and to demonstrate any structural alterations in ZIKV NS3 Helicase loop flexibility subsequent to NITD008 binding.
4. To map out the binding landscape of the ATPase and ssRNA site by demonstrating the chemical characteristics of potent *flavivirus* lead compounds, Lapachol, HMC-HO1 α and Ivermectin at the respective NS3 Helicase binding sites. Insights into the structural and binding features of the ATPase and ssRNA site may be established by:
 - 4.1. Implementing molecular docking to identify structurally favorable molecules from a library of *flavivirus* lead compounds.
 - 4.2. Utilizing the enhanced technique of Accelerated molecular dynamic simulations to validate molecular docking and to assess free-binding energy of the systems by employing the MM/GBSA and per residue decomposition analysis.

1.3 Novelty and Significance of Study

The ZIKV has received considerable attention during 2016. However, prior to the devastating 2015 outbreak in Brazil, the virus was classified as a neglected pathogen similar to Dengue and the West Nile virus (Brasil *et al.* 2016). In recent months, there has been a flood of new discoveries regarding the virus, from evolving modes of viral transmission to viral-linked neurological disorders, unique specificity to host cells and increasing mutation rates (Sironi *et al.* 2016; Cox *et al.* 2016; Pylro *et al.* 2016; Plourde & Bloch 2016; Passi *et al.* 2017; Li *et al.* 2016; Olagnier *et al.* 2016; Bayless *et al.* 2016; Anaya *et al.* 2016).

Fundamental structural and molecular research into the viral targets of ZIKV came into effect in the last year, with the first crystal structure being released in May 2016 (Song *et al.* 2016). The lack in literature regarding the structural dynamics, active binding sites and the minimal understanding of the mechanism of inhibition of ZIKV enzymes prompted us to design a technical route-map toward the design and discovery of potential inhibitors against ZIKV. The study also reported a homology model of the ZIKV NS5 protein, identifying possible binding sites at the MTase and RdRp domains. This was the first account of structure-based virtual screening against the RdRp enzyme, thus assisting scientists from different research domains in designing potential small molecule inhibitors against the viral target.

The scientific community have taken large strides toward developing a effective inhibitor against ZIKV, with preventative clinical trials underway (Marston *et al.* 2016). However, there is still no available FDA approved inhibitor against the virus. We chose to provide insights into the structural dynamics and binding affinities of crucial ZIKV drug targets, being the NS5 MTase and RdRp, as well as the NS3 Helicase. By characterizing the active sites' structural and functional composition, potential small molecule inhibitors may be developed. With the use of CADD techniques, a comprehensive *in silico* perspective is offered to shed light on possible structural characteristics that allow for the inhibition of these enzymes as well as amino acid residues implicated in enzyme activity. Defining the binding landscape will offer prospective design of selective and unique inhibitors with critical pharmacophoric features that will aid in developing targeted and effective small molecule inhibitors.

To this end, the work presented in this thesis is considered to be a fundamental platform in the advancements of research toward targeted drug design/delivery against ZIKV.

References

- Anaya, J.-M. *et al.*, 2016. Zika virus and neurologic autoimmunity: the putative role of gangliosides. *BMC Medicine*, 14(1), pp.1–3.
- Anderson, A.C., 2003. The Process of Structure-Based Drug Design. *Chemistry and Biology*, 10(9), pp.787–797.
- Bayless, N.L. *et al.*, 2016. Zika Virus Infection Induces Cranial Neural Crest Cells to Produce Cytokines at Levels Detrimental for Neurogenesis. *Cell Host & Microbe*, 20, pp.423–428.
- Bollati, M. *et al.*, 2010. Structure and functionality in flavivirus NS-proteins: Perspectives for drug design. *Antiviral Research*, 87(2), pp.125–148.
- Brasil, P. *et al.*, 2016. Zika Virus Outbreak in Rio de Janeiro, Brazil: Clinical Characterization, Epidemiological and Virological Aspects. *PLoS Neglected Tropical Diseases*, 10(4), pp.1–13.
- Brault, J.B. *et al.*, 2016. Comparative Analysis Between Flaviviruses Reveals Specific Neural Stem Cell Tropism for Zika Virus in the Mouse Developing Neocortex. *EBioMedicine*, 10, pp.71–76.
- Broxmeyer, L. & Kanjhan, R., 2016. Does Zika Really Have the Capacity to Affect the Nervous System and Cause Microcephaly or Intracranial Calcifications? *Modern Research in Inflammation*, 5, pp.20–30.
- Chen, B. *et al.*, 2010. MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66(1), pp.12–21.
- Cohen, J., 2016. The race for a Zika vaccine is on. *Infectious Disease*, 351(6273), pp.543–544.
- Cox, B.D., Stanton, R.A. & Schinazi, R.F., 2016. Predicting Zika virus structural biology: Challenges and opportunities for intervention. *Antiviral chemistry & chemotherapy*, 24, pp.118–126.
- D’Ortenzio, E. *et al.*, 2016. Evidence of Sexual Transmission of Zika Virus. *New England Journal of Medicine*, 374(22), pp.2195–2198.
- Foy, B.D. *et al.*, 2011. Probable Non-Vector-borne Transmission of Zika Virus, Colorado, USA. *Emerging Infectious Diseases*, 17(5), pp.880–882.

- Gourinat, A.C. *et al.*, 2015. Detection of zika virus in urine. *Emerging Infectious Diseases*, 21(1), pp.84–86.
- Huang, H.J. *et al.*, 2010. Current developments of computer-aided drug design. *Journal of the Taiwan Institute of Chemical Engineers*, 41(6), pp.623–635.
- Incicco, J.J. *et al.*, 2013. Steady-State NTPase Activity of Dengue Virus NS3: Number of Catalytic Sites, Nucleotide Specificity and Activation by ssRNA. *PLoS ONE*, 8(3),. Epub ahead of print.
- Kollman, P.A. *et al.*, 2016. World Zika Virus Epidemic (2015-16) Monthly Update. *Knoema*, 33.
- Kwong, A.D., Rao, B.G. & Jeang, K.-T., 2005. Viral and cellular RNA helicases as antiviral targets. *Nature reviews. Drug discovery*, 4(10), pp.845–853.
- Li, H. *et al.*, 2016. Zika Virus Infects Neural Progenitors in the Adult Mouse Brain and Alters Proliferation. *Cell Stem Cell*, 19(5), pp.593–598.
- Lu, J. *et al.*, 2012. Computational drug discovery. *Acta Pharmacologica Sinica*, 33, pp.1131–1140. Available at: <http://dx.doi.org/10.1038/aps.2012.109>.
- Mahfuz, M. *et al.*, 2014. Indian Journal of Pharmaceutical and Biological Research (IJPBR) In Silico Modeling and Immunoinformatics Probing Disclose the Epitope Based PeptideVaccine Against Zika Virus Envelope Glycoprotein. *Indian Journal of Pharmaceutical Biological Research*, 2(4), pp.44–57.
- Malone, R.W. *et al.*, 2016. Zika Virus: Medical Countermeasure Development Challenges. *PLoS Neglected Tropical Diseases*, 10(3), pp.1–26.
- Marston, H.. *et al.*, 2016. Considerations for Developing a Zika Virus Vaccine. *The New England journal of medicine*, 375(13), pp.1209–1212.
- Medin, C.L. & Rothman, A.L., 2016. Zika Virus: The Agent and Its Biology, With Relevance to Pathology. *Archives of pathology & laboratory medicine*, 141, pp.33–42.
- Murray, C.L., Jones, C.T. & Rice, C.M., 2009. Architects of Assembly: roles of Flaviviridae nonstructural proteins in virion morphogenesis. *Nat Rev Microbiol*, 6(9), pp.699–708.
- Noble, C.G. *et al.*, 2010. Strategies for development of dengue virus inhibitors. *Antiviral Research*, 85(3), pp.450–462.
- Nowakowski, T.J. *et al.*, 2016. Expression analysis highlights AXL as a candidate zika virus

- entry receptor in neural stem cells. *Cell Stem Cell*, 18(5), pp.591–596.
- Olagnier, D. *et al.*, 2016. Mechanisms of Zika Virus Infection and Neuropathogenesis. *DNA and cell biology*, 35(8), pp.367–372.
- Palomo, A.M., 2016. Zika virus: An international emergency? *Journal of public health policy*, 37(2), pp.133–135.
- Passi, D. *et al.*, 2017. Zika Virus Diseases – The New Face of an Ancient Enemy as Global Public Health Emergency (2016): Brief Review and Recent Updates. *International Journal of Preventive Medicine*, 8(6), pp.1–9.
- Perera, R., Khaliq, M. & Kuhn, R.J., 2008. Closing the door on flaviviruses: Entry as a target for antiviral drug design. *Antiviral Research*, 80(1), pp.11–22.
- Plourde, A.R. & Bloch, E.M., 2016. A Literature Review of Zika Virus. *Emerging Infectious Diseases*, 22(7), pp.1–15.
- Pylro, V.S. *et al.*, 2016. ZIKV – CDB: A Collaborative Database to Guide Research Linking SncRNAs and ZIKA Virus Disease Symptoms. *PLOS Neglected Tropical Diseases*, 10(6), pp.1–9.
- Rasmussen, S.A. *et al.*, 2016. Zika Virus and Birth Defects — Reviewing the Evidence for Causality. *New England Journal of Medicine*, 374, pp.1981–1987.
- Singh, R.K. *et al.*, 2016. Zika Virus - Emergence, evolution, pathology, diagnosis and control: current global scenario and future perspectives - A comprehensive review. *The Veterinary quarterly*, 2176(July), pp.1–43.
- Sironi, M. *et al.*, 2016. Nonstructural Proteins Are Preferential Positive Selection Targets in Zika Virus and Related Flaviviruses. *PLoS Neglected Tropical Diseases*, 10(9), pp.1–18.
- Song, C.M., Lim, S.J. & Tong, J.C., 2009. Recent advances in computer-aided drug design. *Briefings in Bioinformatics*, 10(5), pp.579–591.
- Song, H. *et al.*, 2016. Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. *Nature structural & molecular biology*, 23(5), pp.256–259.
- Tilak, R. *et al.*, 2016. Dengue, chikungunya and the missing entity Zika fever: A new emerging threat. *Medical Journal Armed Forces India*, 72(2), pp.157–163.
- Turmel, M.J. *et al.*, 2016. Late sexual transmission of Zika virus related to. *The Lancet*, 6736(16),

pp.2501–2501.

WHO, 2016. RAPID RISK ASSESSMENT Zika virus disease epidemic : potential association with microcephaly and Guillain-Barré syndrome (1 st update) Main conclusions. *European Centre for Disease Prevention and Control*, (January), pp.1–20.

Zanluca, C., Duarte, C.N. & Santos, D., 2016. Zika virus - an overview. *Microbes and Infection*, 18, pp.295–301.

CHAPTER 2

2. Background on the Zika Virus

2.1 Introduction

The World Health Organization, during 2016, declared ZIKV as a public health emergency due to the virus spreading explosively on a global scale (WHO, 2016). The Zika virus is characterized as an arthropod-borne virus from the *flavivirus* genus and is closely related to the other mosquito-borne viruses including Dengue, Yellow fever and West Nile. As with other *flaviviruses*, the primary vector is the *aedes Aegypti* mosquito found in tropical and sub-tropical areas. The initial symptoms of ZIKV infection in adults were mild influenza-like symptoms that lasted approximately a week. However, as with most viral infections, mutations are eminent, leading to the escalation in virulence and transmission. Scientific communities are in a race to characterize and understand this previously neglected pathogen due to increasing evidence of its responsibility in fetal neurological disorders including microcephaly and Gullian-Barrè syndrome (Ramharack & Soliman 2016).

This chapter contextualizes the ongoing ZIKV research, including the previous outbreaks and the pathogenesis and life cycle of the virus. The structural characteristics of ZIKV will also be reviewed, thus distinguishing possible viral targets in the design of effective and non-toxic therapeutics.

2.2 Epidemiology and Transmission

The Zika virus was first isolated from a pyrexial rhesus monkey in 1947 in Entebbe, Uganda. In 1948, a second isolation was made from the same forest on a group of *Australopithecus africanus* mosquitoes. Due to both these isolations being from the same Zika forest, the virus was labeled as the Zika virus (Dick *et al.* 1952). Although isolations of the virus were analyzed, researchers only detected the virus in humans in 1952 when neutralizing antibodies were picked up in infected sera. Scientists Boorman and Porterfield subsequently studied the transmission of viruses from mosquito to primates and based on further isolations from both mosquito and monkey concluded that mosquitoes acted as vectors for ZIKV (Boorman & Porterfield 1956).

From the 1950s to 2006, ZIKV infection reports were minimal, with sporadic cases in Asian and African countries including Malaysia, Indonesia, Thailand, Pakistan, Nigeria, Senegal, Cameroon and Uganda (Passi *et al.* 2017). The first infectious occurrence of ZIKV was in 2007 in the Yap Islands, Micronesia. Although 80% of the population reported ZIKV symptoms, the virulence was not fatal and there were no hospitalizations (Faye *et al.* 2014; Singh *et al.* 2016; Boeuf *et al.* 2016; Duffy *et al.* 2009). In 2010, a handful of confirmed ZIKV cases were reported in Cambodia, with similar clinical characteristics to the outbreak in the Yap islands (Heang *et al.* 2012). In October 2013, a ZIKV strain analogous to that detected in Cambodia, emerged in French Polynesia with an estimated 19 000 ZIKV infections identified over a two month period (Chen & Hamer 2016; Singh *et al.* 2016).

In May 2015, the ZIKV containing its most virulent strain yet, began its rampage in Brazil. To date, there are close to 1 million cases of ZIKV infection and a third of microcephaly reports in Brazil are linked to perinatal ZIKV transmission (Boeuf *et al.* 2016; Bogoch *et al.* 2016; Lissauer *et al.* 2016). By March 2016, the WHO declared the virus as a public health emergency due to the rapid transmission of the virus to non-endemic regions (WHO 2016a). By November 2016, the virus spread to over 66 countries globally, including Florida, Miami, Singapore, Tonga, Fiji and Cape Verde (Centers for Disease Control 2016).

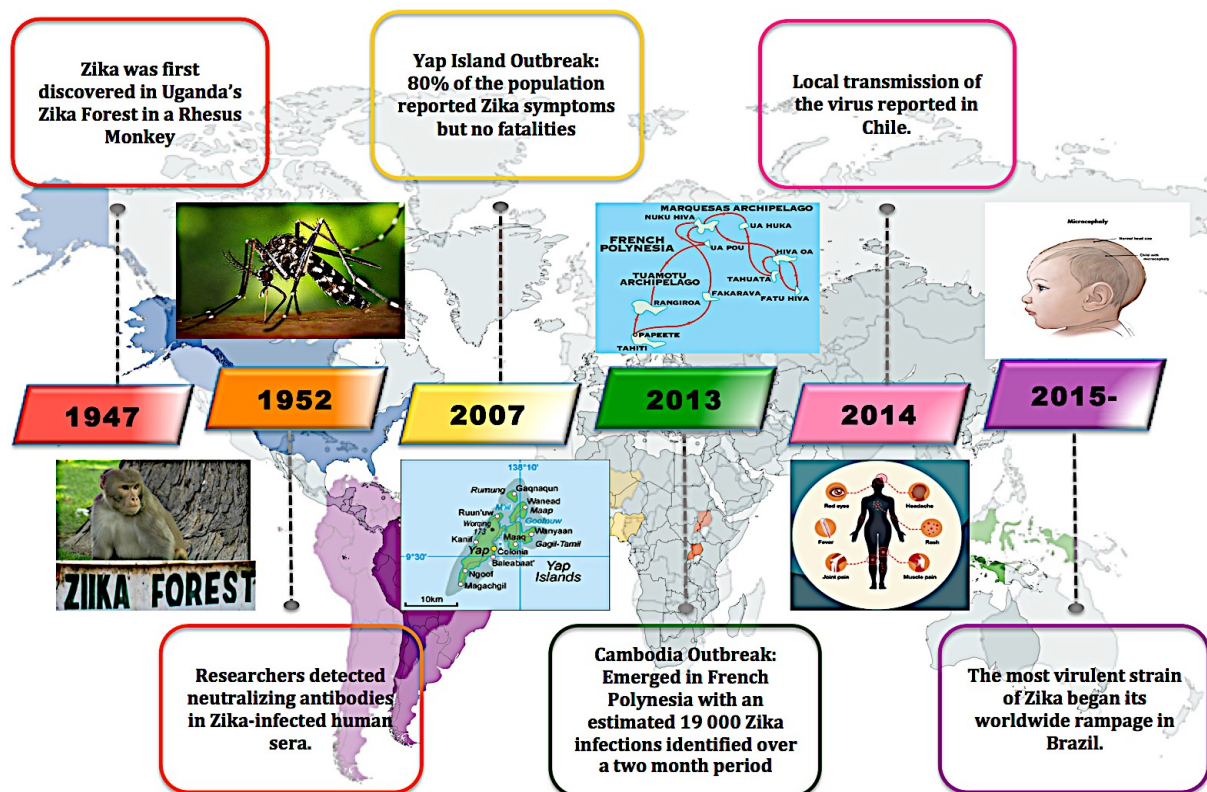


Figure 2.1: Time-line demonstrating the historical outbreaks of ZIKV and the increase in virulence with each strain (Image prepared by author).

The rapid spread of the virus across continents is primarily due to vector transmission via the *Aedes aegypti* or *Aedes albopictus* mosquito. These Vectors are endemic to tropical and sub-tropical, however, due to evolving climates, the mosquitoes have expanded their habitat, thus increasing the number of mosquitoes as vectors of *flaviviruses* (Centers for Disease Control 2016; Shapshak *et al.* 2016). Another reason for the continental dissemination of the virus is the identification of new modes of viral transmission. Recent studies have evidenced ZIKV to be transmitted in a similar fashion to that of the detrimental Human Immunodeficiency Virus (HIV), that is, from mother to fetus via perinatal transmission, blood transfusion and sexual transmission (Hamel *et al.* 2015; Petersen *et al.* 2016).

2.3 Characterizing ZIKV

To date, since the re-emergence of ZIKV and its association with microcephaly and Gullian-Barré syndrome, remarkable efforts have been made in order to provide a better understanding into the

major physiological and molecular mechanisms underlying this infectious disease (WHO 2016b). Fundamental knowledge on the ZIKV structure and life cycle is crucial in designing anti-ZIKV therapeutics as well as exploring the structural implications of ZIKV drug resistance mutations. It is evident by the adverse implications of the Brazilian strain, that the ZIKV is highly mutable, fortifying the challenge of designing efficient inhibitors against the virus. The virus is broadly classified into an East/West African and Asian/Brazilian strains based on sequence resemblance and symptomatic characteristics (Cox *et al.* 2016).

2.3.1 Life Cycle of ZIKV

Subsequent to viral entry into host, the Zika virion attaches to the surface of target cells by interactions between the envelope protein and the host cell surface receptors. The host cell receptors that have been evidenced to mediate virion endocytosis include phosphotydylserine receptor, AXL, as well as DC-SIGN, TIM-1 and Tyro3. Virions undergo this receptor-mediated endocytosis and are internalized to the cell cytoplasm. The viral envelope is then uncoated and the viral RNA is released into the cell cytoplasm. The viral RNA is then translated produce a large polyprotein at the endoplasmic reticulum and is subsequently cleaved into the individual viral proteins, leading to the replication of the viral genome. The viral RNA as well as the structural and non-structural proteins, and some host proteins are involved in the packaging of the viral complex into vesicles and assemble by budding into the endoplasmic reticulum, whereas immature viral particles utilize the host secretory pathway, where virion maturation occurs followed by release from the cell (T. Naga Ravikiran , T. Nagamounika 2016; Brasil *et al.* 2016; Gerold *et al.* 2017; Nugent *et al.* 2016; White 1977; Medin & Rothman 2016). The ZIKV has been evidenced to target a variety of cell types including dendritic cells, human dermal fibroblasts, epidermal keratinocytes and neuronal progenitor cells (Galán-Huerta *et al.* 2016; Boeuf *et al.* 2016).

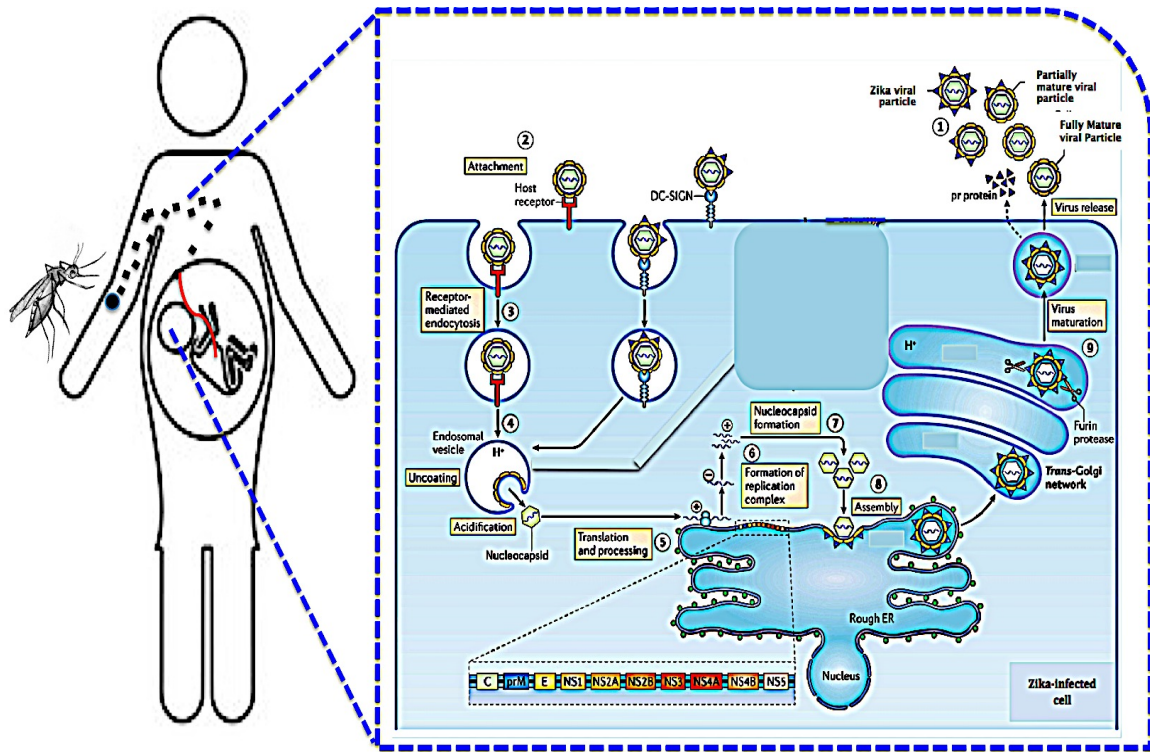


Figure 2.2: Zika Virus Life cycle from vector to host transmission and replication (Adapted from (Screaton *et al.* 2015)).

2.3.2 Structural Characteristics of ZIKV

The ZIKV is an enveloped icosahedral virus that is made up of a single-stranded, positive-sense genome. The enveloped virion comprises of an 11 kilobase genome consisting of 10,794 nucleotides encoding 3,419 amino acids (Hayes 2009). The open reading frame (ORF) of the 5' and 3' untranslated region (UTR) encodes a polyprotein that is cleaved into three structural proteins being the capsid, precursor membrane, and envelope. Seven non-structural (NS) proteins are also found in this assembly, namely, NS1, NS2a, NS2b, NS3, NS4a, NS4b, and NS5 (largest viral protein) (Haddow *et al.* 2012; Boeuf *et al.* 2016), in which the genomic protein organization is 5'-C-prM-E-NS1-NS2a-NS2b-NS3-NS4a-NS4b-NS5-3' (White *et al.* 2016). The genomic RNA of ZIKV contains an m⁷gpppAmpN₂ at the 5' end and lacks a poly-A tail at the 3' end (White *et al.* 2016; Cox *et al.* 2016). There is also a highly conserved 90-120-nucleotide strand near the 3' end that develops into a hairpin loop that is crucial for replication (Passi *et al.* 2017; Mumtaz *et al.* 2016). Of the non-structural proteins, NS1, NS3 and NS5 are highly conserved

whereas the NS2a, NS2b, NS4a and NS4b are small and hydrophobic (Galán-Huerta *et al.* 2016). Of critical importance is the proteolytic cleavage of prM to give the pr and M protein, which is produced by furin-like protease located in the trans-Golgi network during the egress of the particles and promote the maturation of the virions (Saiz *et al.* 2016).

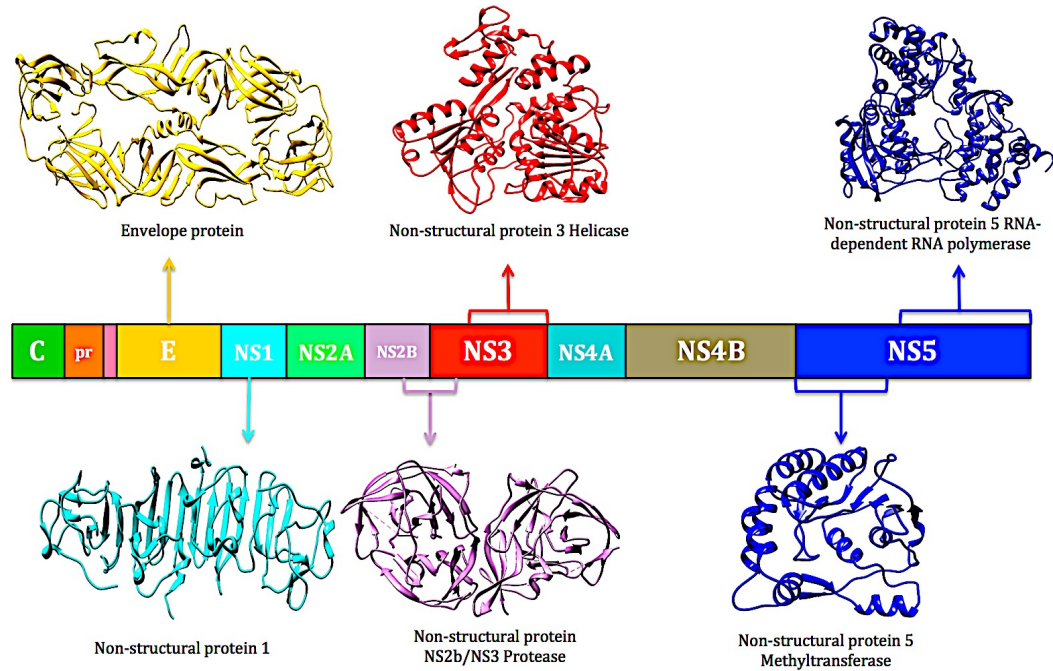


Figure 2.3: Cleaved ZIKV polyprotein demonstrating available protein crystal structures (PDB codes: 5IY3, 5JHM, 5JMT, 5T1V, 5KQR, 5U04) (Prepared by Author).

2.3.3 ZIKV Pathogenesis and Clinical Features

Although isolated in the early 1900s, many of the distinguishing clinical and pathogenic features have only been discovered in recent years. Initially, the clinical characteristics of the virus were minor, consisting of flu-like symptoms including swollen lymph nodes, maculopapular skin rashes and joint pains (Mahfuz *et al.* 2015; Singh *et al.* 2016; Plourde & Bloch 2016). Current research has now associated the virus with multiple-organ failure and thrombocytopenic purpura (Miner & Diamond 2016). The virus has also been evidenced to cause uveitis, a inflammatory eye disease in adults (Furtando *et al.*, 2016) and conjunctivitis in approximately 15% of patients (Miner & Diamond 2016). The most detrimental complication surrounding ZIKV infection is its ability to target neural progenitor cells, thus leading to fetal central nervous system disorders such

as microcephaly and cerebral calcification as well as Gullian-Barré syndrome which has been described in adult infection.

After the local replication in host cells, ZIKV is distributed to the heart, muscle and central nervous system, as well as across the placental barrier to the fetus (Singh *et al.* 2016). Once in the amniotic fluid, ZIKV has been shown to infect neural progenitor cells, thus triggering apoptosis (programmed cell death) of the host cell (Tang *et al.* 2016). This could be one of the potential mechanisms by which ZIKV causes fetal microcephaly (El Costa *et al.* 2016; Huang *et al.* 2016; Ghouzzi *et al.* 2017; White *et al.* 2016). Gullian-Barré is caused by the demyelination of nerves leading to muscle weakness, tingling in arms and legs and in severe cases, paralysis. The ZIKV may be able to infect myelin directly or via autoimmune-mediated targeting of neurons and glial cells (Miner & Diamond 2016).

A study by Grant *et al.*, 2016, demonstrated the degradation of STAT2 by ZIKV NS5 protein, thus inhibiting immune-response cells IFN-1 and the innate immune response. The characteristic features of ZIKV, being its ability to pass through the blood-brain-barrier and placental barrier, as well as target neuronal cells and dampen the host immune response, allow for its persistence and replication in the human host (Grant *et al.* 2016). Targeting specific key proteins of the virus as well as possibly targeting invaded host machinery will allow for the inhibition of the virus, halting the progression of any downstream complications.

2.4 Rationale of ZIKV Enzymes as Potential Therapeutic Targets

Hughes *et al.* (2010) stated that the potential of a protein as a therapeutic target and its effectiveness in drug design is essential for determining the biological utility of the protein (Hughes *et al.* 2011). ZIKV contains a plethora of viral proteins that may act as targets in drug design. To identify inhibitors that specifically halt essential steps in the ZIKV life cycle, fundamental characteristics of each protein need to be established and essential proteins need to be identified.

The ZIKV is composed of an inner shell formed by interacting subunits of the capsid (C) which is able to interact with genomic RNA, an intermediate shell composed of the membrane (M) and an

outer shell containing the viral glycoproteins (E and prM). Disruption of the structural proteins by antibodies or small molecule inhibitors may interfere with structural protein interactions thereby inhibiting virion assembly and capsid dimerization (Cox *et al.* 2016; Ekins *et al.* 2016; Sironi *et al.* 2016).

Replication of viral RNA requires the activities of several non-structural proteins as well as utilization of specific host proteins. The NS1 and NS2a proteins have shown to evade the innate immune system by acting as antagonists against Toll-like receptor-3 (TLR3) and interferon (IFN) α/β , thus providing a biochemical pathway as a starting point in the design of antivirals against these NS proteins (Geiss *et al.* 2010). The NS4a/b proteins also prove to be potential targets for therapeutic intervention as a study by Liang *et al.* (2016) has evidenced the NS proteins to inhibit the AKT/mTOR pathway, thus halting neurogenesis and inducing autophagy (Liang *et al.* 2016). Of the non-structural proteins however, the NS3 and NS5 proteins are considered as prime targets for antiviral development due to their essential roles in ZIKV RNA replication:

2.4.1 NS5 Protein

The largest non-structural protein translated from the ZIKV genome, with a molecular weight of approximately 103 kDa, is the NS5 protein (Figure 2.4) (Cox *et al.* 2016). It is comprised of a Methyltransferase (MTase) N-terminal RNA capping domain and a C-terminal domain with RNA-dependent RNA polymerase (RdRp) enzymatic activity. The ZIKV genome sustains a 5' cap that is methylated to facilitate stability and evasion of host immune responses. The RdRp domain of the NS5 is crucial for RNA replication as it initiates RNA synthesis by generating negative-sense RNA from a positive-strand template. The synthesized strand then facilitates the generation of a positive-stranded RNA during viral replication (Alshiraihi *et al.* 2016). The 5' end of the viral RNA molecule includes a methylated cap comprising of a guanine nucleotide tethered to the first nucleotide of the RNA. Like all polymerases, the structure of ZIKV RdRp portrays a right hand with characteristic fingers, palm, and thumb subdomains. There are two cavities located in the thumb subdomain; however, there is no biological relevance of the cavities to date (Alshiraihi *et al.* 2016; Malet *et al.* 2008 Zou *et al.* 2011) As mentioned above, there are two strains of ZIKV, being the African and Asian/Brazilian Strains, The substitutions between African and Asian strains occur mostly on the surface of the RdRp domain. The K/R280N,

H449Q, and G587K ZIKV mutations occur in the finger region and two mutations are found in the thumb domain (A784S and D867N) (Cox *et al.* 2016).

The MTase domain of NS5 is a 33 kDa protein comprising of 1-260 amino acids. There are multiple active-binding sites, including a positively charged RNA binding site, a site for the methyl donor S-Adenosyl Methionine (SAM), and a Guanine Triphosphate (GTP) cap-binding pocket (Alshiraihi *et al.* 2016; De Oliveira *et al.* 2014). The core domain contains four α -helices surrounding a seven-stranded β -sheet. The N-terminal segment comprises a helix-turn-helix motif followed by a β -strand and an α -helix. The C-terminal region consists of an α -helix and two β -strands. The functional domain of the MTase is found at the N-terminal region of the protein and allows for the methylation of both the N-7 position of the 5' guanine cap as well as the ribose 2'-OH position of the first transcribed nucleotide. SAM methionine interacts with S56, D146, G86 and W87, whereas, D146 is integral to a motif that is essential for N7 and 2'O methylation (Bollati *et al.* 2010; Sampath & Padmanabhan 2009; Cox *et al.* 2016; Zou *et al.* 2011).

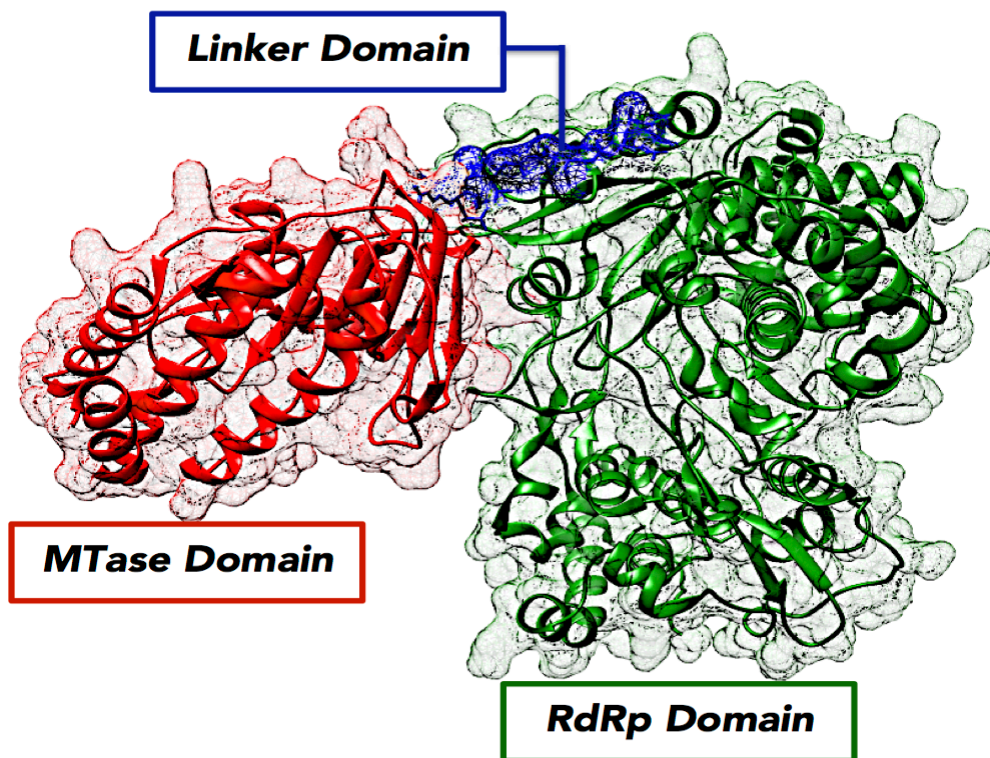


Figure 2.4: Crystal structure of ZIKV NS5 protein. The protein comprises of three domains, the N-terminal Mtase domain (residues 1-262) (red), the C-terminal RdRp domain (residues 273-907) (green) and linker domain (residues 263-272) (blue), (PDB code: 5TFR) (Prepared by Author).

2.4.2 NS3 Protein

The NS3 protein consists of 618 amino acids and a serine-protease domain at its N-terminal and an ATP-driven Helicase domain and RNA triphosphate at its C-terminal. The protein is also involved in viral assembly independently of the known enzymatic activity mentioned (Lescar *et al.* 2008). During the ZIKV life cycle, the NS3 protein directly interacts with the NS5 polymerase to effectively multiply the viral genome. Studies have shown that impairment of either domain on either protein lead to non-infectious production of the viral particles.

The catalytic triad, Ser¹³⁵-His⁵¹-Asp⁷, NS2b/NS3 protease is 375 kDA protein that is dependent on the association of the 14 kDA, 40 amino acid NS2b cofactor, for its activity (Bollati *et al.* 2010). The two NS proteins are covalently linked via a Gly₄-Ser-Gly₄ sequence, displaying strong peptidolytic activity (Lei *et al.* 2016). One unique feature of the ZIKV protease is its quasi-twofold dimer symmetry. In the dimer, the substrate-binding sites of the two monomers along with the bound inhibitor face each other. The dimer has an opening on both sides that allows for the substrate to be accessed from both active sites (Lei *et al.* 2016). The protease functions by cleaving the polypeptide chain between bonds NS2a-NS2b, NS2b-NS3, NS3-NS4a and NS4b-NS5. This cleavage is essential for viral replication as the activities of the NS proteins are dependent on their cleavage at precise amino acids (Chen *et al.* 2016; White *et al.* 2016).

The ZIKV helicase comes from the superfamily helicases, SF2 and is found at the C-terminal of the NS3 protein and requires an ATP-driven molecular motor. The structural characteristics of the ZIKV NS3 helicase consists of three domains of approximately 440-450 residues: domain I (residues 182-327), domain II (residues 328-480) and domain III (residues 481-617), as well as a P-Loop (residues 196-203) which is located at the ATP-binding site of domain I (Jain *et al.* 2016; Hongliang Tian *et al.* 2016) (Figure 2.5). The stimulation by RNA allows the helicase domain to exhibit intrinsic nucleoside triphosphatase activity, which then allows for the unwinding of viral RNA to facilitate replication of the viral genome with the NS5 RdRp (H Tian *et al.* 2016). The inhibition of either one of the binding sites, the RNA-binding groove or the ATP-binding site, leads to insufficient viral replication and maturation (Sampath & Padmanabhan 2009).

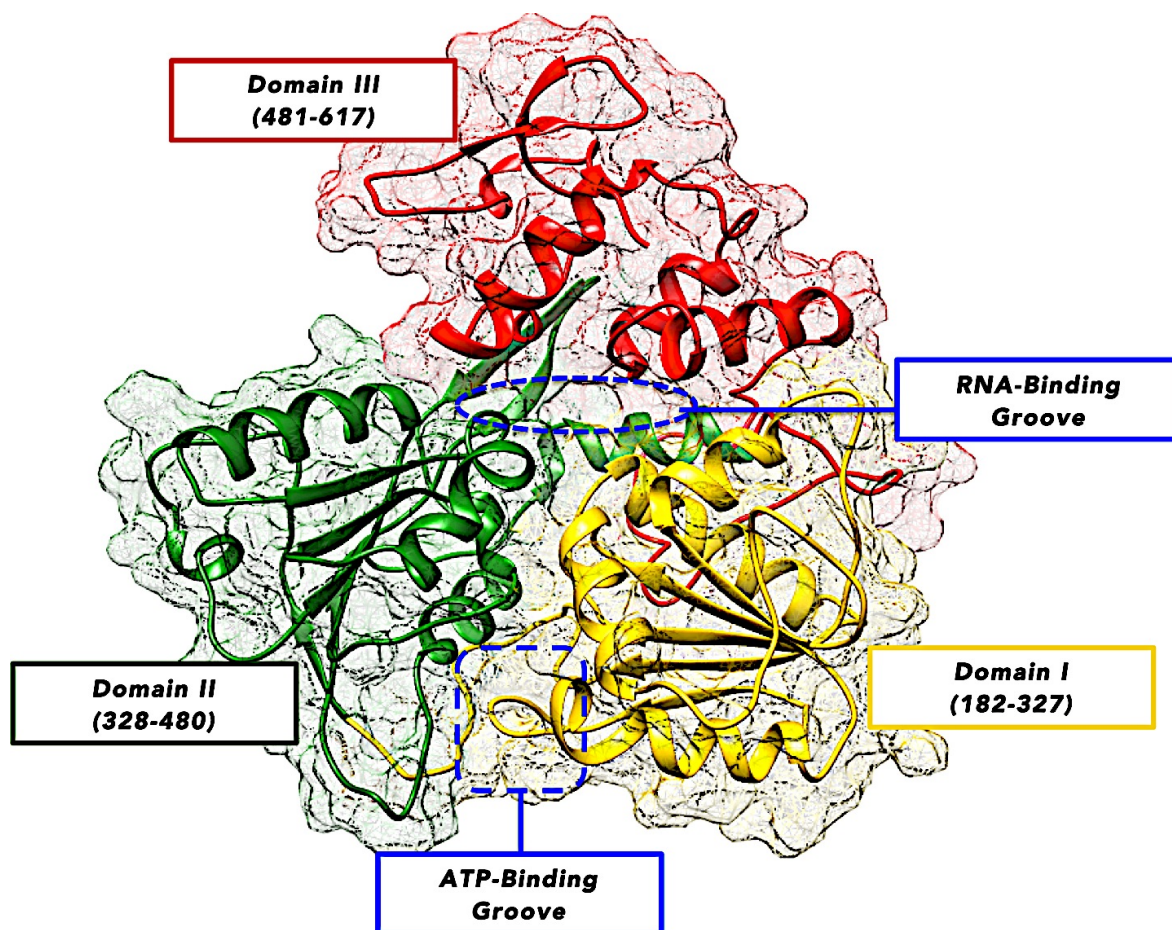


Figure 2.5: Crystal structure of ZIKV NS3 Helicase Protein (PDB code: 5GJC) depicting the three domains and two-active binding regions (blue) that form hydrophobic pockets for ATP and ssRNA binding (Prepared by Author).

2.4.3 Three-dimensional structures of ZIKV NS5 and NS3 proteins

Prior to 2016, there were no available crystal structures of any of the ZIKV proteins. However, there has been a flood of scientific knowledge released in the past two years regarding the fundamental characteristics of ZIKV and the basis for ZIKV rational drug design. This has allowed for the release of crystal structures of ZIKV proteins, providing new insights on the structural features of these targets. Table 1 summarizes the currently available PDB-deposited crystal structures of ZIKV NS5 and NS3 proteins.

Table 2.1: Overview of the currently PDB-deposited Crystal Structures of ZIKV NS5 and NS3 Proteins.

NS Protein	Resolution (Å)	Ligand/ Inhibitor	PDB Entry	Reference, Date of Publication
Zika Virus NS5 RNA-dependent RNA polymerase	2.31	ZN	5TIT	Godoy <i>et al</i> (To be Published)
Zika Virus NS5 RNA-dependent RNA polymerase	1.9	PO4, ZN	5UO4	Godoy <i>et al</i> (To be Published)
Structure of Zika virus NS5	3.28	<u>GOL, SAH, S</u> <u>O4, ZN</u>	5TMH	Wang <i>et al</i> (To be Published)
Zika virus NS5 methyltransferase	2.01	CL, GOL, SAM, SO4	5M5B	Coutard <i>et al</i> (March 2017)
Zika virus NS5 Methyltransferase in complex with GTP and SAH	2.05	GTP, PO4, POP, SAH, SIN	5GOZ	Zhang <i>et al</i> (November 2016)
Zika virus NS5 Methyltransferase in complex with GTP and SAH	2.44	GTA, NI, SAH, SO4	5GP1	Zhang <i>et al</i> (November 2016)
Zika Virus NS5 Protein	3.05	SAH, ZN	5TFR	Longnecker <i>et al</i> (To be Published)
NS5 methyltransferase from Zika virus bound to S-adenosylmethionine	1.33	CL, PO4, SAM	5KQR	Coloma <i>et al</i> (September 2016)
Zika virus bound to S-adenosylmethionine and 7-methylguanosine-5'-diphosphate	1.5	ACT, GOL, M7G, PO4, SAM	5KQS	Coloma <i>et al</i> (September 2016)
Zika NS3 helicase:RNA complex	1.6	ACT, FLC	5MF X	Jenkins <i>et al</i> (To be Published)
Unlinked NS2b-NS3 Protease from Zika Virus and its complex with a Reverse Peptide Inhibitor	1.58	-	5GPI	Zhang <i>et al</i> (December 2016)
Unlinked NS2b-NS3 Protease from Zika Virus in complex with a compound fragment	2.0	7HQ, ACT	5H4I	Zhang <i>et al</i> (December 2016)
Zika virus NS3 helicase	2.05	K, TRS	5TXG	Nocadello <i>et al</i> (To be Published)
Apo structure	1.4	EDO	5JWH	Cao <i>et al</i> (To be Published)
Apo structure	1.69	ATP, CL, MN	5K8I	Cao <i>et al</i> (To be Published)

Apo structure	1.75	CL, EDO, GSP, MPD	5K8L	Cao <i>et al</i> (To be Published)
ZIKV NS3 helicase in complex with GTP-gamma S and a magnesium ion	1.85	CL, GSP, MG	5K8T	Cao <i>et al</i> (To be Published)
Apo structure	1.6	ADP, CL, EDO, MN	5K8U	Cao <i>et al</i> (To be Published)
NS2b-NS3 Protease from Zika Virus caught after self-cleavage	1.84	CL	5GJ4	Phoo <i>et al</i> (November 2016)
Zika virus NS2b-NS3 protease in Apo form	3.1	-	5T1V	Nocadello <i>et al</i> (To be Published)
Zika virus NS3 helicase in complex with ssRNA	1.7	-	5GJB	Tian <i>et al</i> (August 2016)
Zika virus NS3 helicase in complex with ATP	2.2	ATP, MN	5GJC	Tian <i>et al</i> (August 2016)
NS3 Helicase from the French Polynesia strain of the Zika virus	1.62	ACT, POP	5JRZ	Jain <i>et al</i> (August 2016)
Zika virus NS2b-NS3 protease in complex with a boronate inhibitor	2.7	6T8	5LC0	Lei <i>et al</i> (July 2016)
Zika virus NS3 helicase	1.8	-	5JMT	Tian <i>et al</i> (June 2016)

2.5 Targeting Host Proteins in ZIKV Therapy

During the ZIKV replication cycle, host cell machinery is imperative in the translation of viral RNA and maturation of the replicated virus, thus targeting host proteins and pathways may be key to effective inhibition of viral replication.

One of the most researched host proteins in flavivirus infection is the endoplasmic reticulum glucosidase. These proteins allow for the cleavage of the terminal glucose from the glycan found at the glycosylation-site of the prM and envelope protein, thus leading to its maturation of the envelope protein (Stahla-Beek *et al.* 2012). Studies have shown that many flaviviruses, including ZIKV, have a N-glycosylation at Asn154 (Ekins *et al.* 2016; Sirohi *et al.* 2016). Castanospermine (CST) and deoxynojirimycin (DNJ) have been established as potent inhibitors of alpha-glucosidases, thus preventing the early stages of glycosylation (Courageot *et al.* 2000).

Hamel *et al* (2015) described the importance of dendritic cell-specific intracellular adhesion molecule 3- grabbing non-integrin (DC-SIGN), TIM-1 and TAM receptors in the attachment and entry of ZIKV into the host cell before replication can occur (Hamel *et al.* 2015). Small interfering RNA (siRNA) was also shown to completely inhibit the expression of the above proteins after 48 hours. Other informative publications on siRNA inhibition of flavivirus host machinery include a review by Hirsch (2010), an *in silico* based experimental study on Dengue virus by Noppakunmongkolchai *et al* (2016) and the silencing of the 3' UTR of ZIKV genome by Shawan *et al* (2015) (Mahfuz *et al.* 2015; Noppakunmongkolchai *et al.* 2016; Hirsch 2010).

A recent study published in June, 2016 by *Nature* identified host endoplasmic reticulum-associated signal peptidase complex (SPCS) to be necessary for the proper cleavage of ZIKV prM and envelope proteins. The authors also demonstrated that the loss of SPCS signaling leads to a dramatic decrease in Dengue, Yellow fever, West Nile, JEV and Hepatitis C viruses (R. Zhang *et al.* 2016). Nowakowski *et al* (2016) also found membrane receptor AXL to have potential as a host target as it facilitates the entry of ZIKV into the host cell (Nowakowski *et al.* 2016). Wells *et al* (2016) however rejected this theory as he demonstrated AXL-knockout to still allow for ZIKV entry. He proposed an attachment factor, TYRO3 to be a possible host target (Wells *et al.* 2016).

2.6 The Scientific Advancements of ZIKV Anti-viral Therapy

The Asian/Brazilian strain of ZIKV has already been associated with irreversible chronic central nervous system (CNS) conditions as mentioned above. The concerns of the scientific and clinical community are the consequences of Zika viral mutations, thus suggesting the urgent need for viral inhibitors. There have been large strides in vaccine development against the virus but there are still no licensed treatments available. Rapid rational drug design and discovery research is fundamental in the production of potent inhibitors against the virus that will not just mask the virus, but destroy it completely. Recent research has found that one of the characteristic features of ZIKV is that it targets neuronal cells (Millichap 2016; Miner & Diamond 2016; Mlakar *et al.* 2016; Tang *et al.* 2016). Consequently, any new drugs that may be discovered will have to pass through the blood-brain-barrier. Currently, there are number of promising prevention therapies and potential treatment options including small molecules (some of which have previously been approved by FDA to treat other diseases), vaccine candidates, and neutralizing purified antibodies still being tested. Below are an overview of such experimental therapies:

2.6.1 Preventative Antibodies and Vaccines

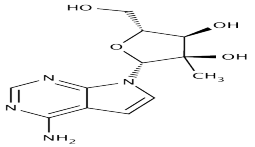
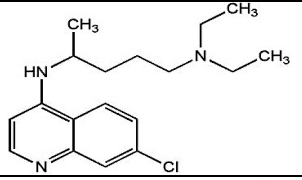
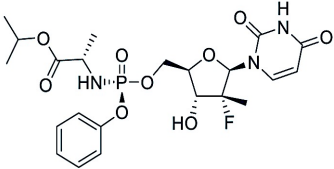
“Prevention is better than cure”, a quote that is true to its meaning. Vaccination is one of the most effective forms of protection against a viral infection. Immunization with an inactivated vaccine will be the most secure route with ZIKV infection as it will be safe to use by pregnant woman (Cohen 2016). Marston *et al.*, 2016, recently published a set of considerations for developing a ZIKV vaccine that will allow for safe and effective control of the virus based on focused planning and evaluation (Marston *et al.* 2016). Mahfuz *et al.*, 2014, began the design of epitope-based vaccines against ZIKV envelope glycoprotein; however, this was an introductory approach and was not validated in subsequent studies (Mahfuz *et al.* 2014). In July 2016, a study was done on repurposing Dengue virus antibodies as inhibitors of ZIKV at different pH levels. Results showed CryoEM structures of potent *flavivirus* antibody C10 bound to ZIKV envelope protein at pH 6.5 and pH5.0, suggesting a new candidate in ZIKV vaccine therapeutics (S. Zhang *et al.* 2016). Another monoclonal antibody was identified to bind to the glycan loop of the envelope protein, thus potentially inhibiting the binding of ZIKV to host cell receptors (Barba-Spaeth *et al.* 2016). Abbink *et al.*, 2016, showed promising results of a purified inactivated viral vaccine, which induced ZIKV-specific neutralizing antibodies and immunized a test group of Rhesus monkeys.

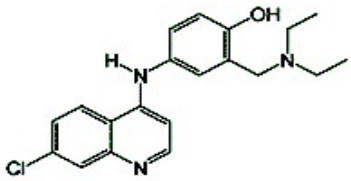
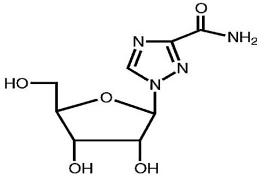
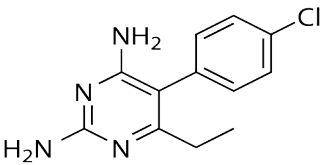
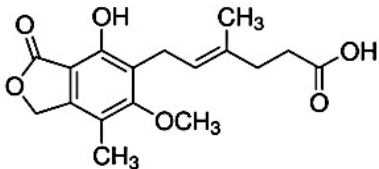
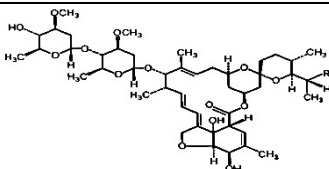
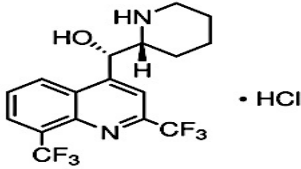
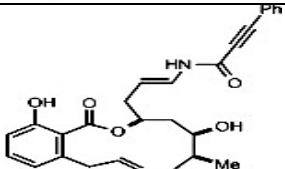
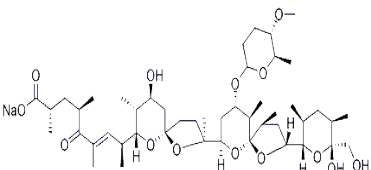
A plasmid DNA vaccine and a single-shot recombinant rhesus adenovirus serotype 52 vector expressing ZIKV prM-E also produced neutralizing antibodies and completely protected monkeys against ZIKV infection. These results are promising and clinical trials are currently underway with hopes of a FDA approved vaccine (Abbink *et al.* 2016; Larocca *et al.* 2016).

2.6.2 Small Molecule inhibitors

Antibodies and vaccines are often expensive and require specific conditions for the transport and storage of the vaccines. This proves to be problematic in developing countries where funds and facilities are limited. Development of small molecule inhibitors cost less, are produced faster, they are stored large quantities and are generally more accessible. Due to the rapid spread of ZIKV infection on a global scale and the detrimental long-term complications, the scientific community has turned to, rather than designing and synthesizing new drugs, to ‘repurpose’ *flavivirus* FDA approved drugs for ZIKV (Mumtaz *et al.* 2016; Wahid *et al.* 2016). Table 2 summarizes the ZIKV drug candidates based on related viral inhibitors.

Table 2.2: Most Popular Repurposed Drugs as ZIKV Inhibitors

DRUG	MECHANISM OF ACTION	STRUCTURE
2-C-methyladenosine	Inhibition of RdRp (Eyer <i>et al.</i> 2016).	
Chloroquine	Endocytosis blocking agent (Delvecchio <i>et al.</i> 2016).	
Sofosbuvir	Hepatitis C RdRp inhibitor (Kryger <i>et al.</i> 2013).	

Amodiaquine	Antimalarial drug (Hinton <i>et al.</i> 2016).	
Ribavirin	Inhibition of RdRp (Kryger <i>et al.</i> 2013).	
Pyrimethamine	Dihydrofolate reductase antagonist (Barrows <i>et al.</i> 2016).	
Mycophenolic Acid	Inosine- 5'-monophosphate dehydrogenase inhibitor (Barrows <i>et al.</i> 2016).	
Ivermectin	Inhibition of RNA-binding (Barrows <i>et al.</i> 2016).	
Mefloquine HCl	Autophagy/ disrupts lysosomal pH (Barrows <i>et al.</i> 2016).	
Saliphenylhalamide	Viral RNA production inhibitors (Kuivanam 2016)	
Nanchangmycin	Viral entry inhibitor (Bausch)	

References

- Abbink, P. *et al.*, 2016. Protective efficacy of multiple vaccine platforms against Zika virus challenge in rhesus monkeys. *Science*, 353(6304), pp.1129–1132.
- Alshiraihi, I.M.A. *et al.*, 2016. Targeting the NS5 Protein of Zika Virus. *Journal of Multidisciplinary Engineering Science Studies*, 2(12), pp.1237–1240.
- Barba-Spaeth, G. *et al.*, 2016. Structural basis of potent Zika–dengue virus antibody cross-neutralization. *Nature*, 536, pp.48–53.
- Barrows, N.J. *et al.*, 2016. A Screen of FDA-Approved Drugs for Inhibitors of Zika Virus Infection. *Cell Host and Microbe*, 20(2), pp.259–270.
- Boeuf, P. *et al.*, 2016. The global threat of Zika virus to pregnancy : epidemiology , clinical perspectives , mechanisms , and impact. *BMC Medicine*, 14(112), pp.1–9.
- Bogoch, I.I. *et al.*, 2016. Anticipating the international spread of Zika virus from Brazil. *The Lancet*, 387(10016), pp.335–336.
- Bollati, M. *et al.*, 2010. Structure and functionality in flavivirus NS-proteins: Perspectives for drug design. *Antiviral Research*, 87(2), pp.125–148.
- Boorman, J.P.T. & Porterfield, J.S., 1956. A simple technique for infection of mosquitoes with viruses transmission of Zika virus. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 50(3), pp.238–242.
- Brasil, P. *et al.*, 2016. Zika Virus Outbreak in Rio de Janeiro, Brazil: Clinical Characterization, Epidemiological and Virological Aspects. *PLoS Neglected Tropical Diseases*, 10(4), pp.1–13.
- Centers for Disease Control, 2016. Zika Virus. *Centers for Disease Control and Prevention: Zika Virus Home*, pp.1–12.
- Chen, L. & Hamer, D.H., 2016. Zika Virus: Rapid Spread in the Western Hemisphere. *Annals of Internal Medicine*, (February), pp.1–3.
- Chen, X. *et al.*, 2016. Mechanisms of activation and inhibition of Zika virus NS2B-NS3 protease. *Cell Research*, 26(11), pp.1260–1263.
- Cohen, J., 2016. The race for a Zika vaccine is on. *Infectious Disease*, 351(6273), pp.543–544.

- El Costa, H. *et al.*, 2016. ZIKA virus reveals broad tissue and cell tropism during the first trimester of pregnancy. *Scientific Reports*, 6(1), p.35296.
- Courageot, M. *et al.*, 2000. α -Glucosidase Inhibitors Reduce Dengue Virus Production by Affecting the Initial Steps of Virion Morphogenesis in the Endoplasmic Reticulum α - Glucosidase Inhibitors Reduce Dengue Virus Production by Affecting the Initial Steps of Virion Morphogenesis in. *Journal of Virology*, 74(1), pp.564–572.
- Cox, B.D., Stanton, R.A. & Schinazi, R.F., 2016. Predicting Zika virus structural biology: Challenges and opportunities for intervention. *Antiviral chemistry & chemotherapy*, 24, pp.118–126.
- Delvecchio, R. *et al.*, 2016. Chloroquine inhibits Zika Virus infection in different cellular models. *bioRxiv*, p.51268.
- Dick, G., Kitchen, S.F. & Haddow, A.J., 1952. Zika virus. I. Isolations and serological specificity. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 46(5), pp.509–520.
- Duffy, M.R. *et al.*, 2009. Zika virus outbreak on Yap Island, Federated States of Micronesia. *The New England journal of medicine*, 360(24), pp.2536–2543.
- Ekins, S. *et al.*, 2016. Open drug discovery for the Zika virus. *F1000Research*, 5(150), pp.1–14.
- Eyer, L. *et al.*, 2016. Nucleoside inhibitors of Zika virus. *Journal of Infectious Diseases*, 27, pp.1–13.
- Faye, O. *et al.*, 2014. Molecular Evolution of Zika Virus during Its Emergence in the 20th Century. *PLoS Neglected Tropical Diseases*, 8(1), pp.1–10.
- Galán-Huerta, K.A. *et al.*, 2016. The Zika virus disease: An overview. *Medicina Universitaria*, 18(71), pp.115–124.
- Geiss, B.J. *et al.*, 2010. Focus on flaviviruses : current and future drug targets. *Future medicinal Chemistry*, 1(2), pp.1–28.
- Gerold, G. *et al.*, 2017. Protein Interactions during the Flavivirus and Hepacivirus Life Cycle. *Molecular & Cellular Proteomics*, 16, pp.S75–S91.
- Ghouzzi, V. El *et al.*, 2017. ZIKA virus elicits P53 activation and genotoxic stress in human neural progenitors similar to mutations involved in severe forms of genetic microcephaly

- and p53. *Cell Death and Disease*, 8(1), pp.1–10.
- Grant, A. *et al.*, 2016. Zika Virus Targets Human STAT2 to Inhibit Type I Interferon Signaling. *Cell Host and Microbe*, 19(6), pp.882–890.
- Haddow, A.D. *et al.*, 2012. Genetic characterization of Zika virus strains: Geographic expansion of the Asian lineage. *PLoS Neglected Tropical Diseases*, 6(2), p.doi: 10.1371/journal.pntd.0001477, Epub ahead of print.
- Hamel, R. *et al.*, 2015. Biology of Zika Virus Infection in Human Skin Cells. *Journal of virology*, 89(17), pp.8880–96.
- Hayes, E.B., 2009. Zika virus outside Africa. *Emerging Infectious Diseases*, 15(9), pp.1347–1350.
- Heang, V. *et al.*, 2012. Zika virus infection, Cambodia, 2010. *Emerging Infectious Diseases*, 18(2), pp.349–351.
- Hinton, T.M. *et al.*, 2016. Polyanionic Macromolecular Prodrugs of Ribavirin: Antiviral Agents with a Broad Spectrum of Activity. *Advanced Healthcare Materials*, 5(5), pp.534–540.
- Hirsch, A.J., 2010. in Viral Replication. *Future Medicinal Chemistry*, 5(2), pp.303–311.
- Huang, W.-C. *et al.*, 2016. Zika virus infection during the period of maximal brain growth causes microcephaly and corticospinal neuron apoptosis in wild type mice. *Scientific Reports*, 6(34793), pp.1–8.
- Hughes, J.P. *et al.*, 2011. Principles of early drug discovery. *British Journal of Pharmacology*, 162(6), pp.1239–1249.
- Jain, R. *et al.*, 2016. Structure of the NS3 helicase from Zika virus. *Nature structural & molecular biology*, 2(July), pp.1–4.
- Kryger, M.B.L. *et al.*, 2013. Macromolecular prodrugs of ribavirin combat side effects and toxicity with no loss of activity of the drug. *Chemical communications*, 49(26), pp.2643–5.
- Larocca, R.A. *et al.*, 2016. Vaccine protection against Zika virus from Brazil. *Nature*, pp.1–19.
- Lei, J. *et al.*, 2016. Crystal structure of Zika virus NS2B-NS3 protease in complex with a boronate inhibitor. *Science*, 353(6298), pp.503–5.
- Lescar, J. *et al.*, 2008. Towards the design of antiviral inhibitors against flaviviruses: The case for

- the multifunctional NS3 protein from Dengue virus as a target. *Antiviral Research*, 80(2), pp.94–101.
- Liang, Q. *et al.*, 2016. Zika Virus NS4A and NS4B Proteins Deregulate Akt-mTOR Signaling in Human Fetal Neural Stem Cells to Inhibit Neurogenesis and Induce Autophagy. *Cell Stem Cell*, 19(5), pp.663–671.
- Lissauer, D., Smit, E. & Kilby, M.D., 2016. Zika virus and pregnancy. *An International Journal of Obstetrics and Gynaecology*, 123(8), pp.1258–1263.
- Mahfuz, M. *et al.*, 2015. Design and Prediction of Potential RNAi (siRNA) Molecules for 3' UTR PTGS of different strains of Zika Virus: A Computational Approach. *Nature and Science*, 1(2), pp.37–50.
- Mahfuz, M. *et al.*, 2014. Indian Journal of Pharmaceutical and Biological Research (IJPBR) In Silico Modeling and Immunoinformatics Probing Disclose the Epitope Based Peptide Vaccine Against Zika Virus Envelope Glycoprotein. *Indian Journal of Pharmaceutical Biological Research*, 2(4), pp.44–57.
- Malet, H. *et al.*, 2008. The flavivirus polymerase as a target for drug discovery. *Antiviral Research*, 80(1), pp.23–35.
- Marston, H. *et al.*, 2016. Considerations for Developing a Zika Virus Vaccine. *The New England journal of medicine*, 375(13), pp.1209–1212.
- Medin, C.L. & Rothman, A.L., 2016. Zika Virus: The Agent and Its Biology, With Relevance to Pathology. *Archives of pathology & laboratory medicine*, 141, pp.33–42.
- Millichap, J.G., 2016. Zika Virus Infection and Microcephaly. *Pediatric Neurology Briefs*, 30(1), pp.8–8.
- Miner, J.J. & Diamond, M.S., 2016. Understanding how zika virus enters and infects neural target cells. *Cell Stem Cell*, 18(5), pp.559–560.
- Mlakar, J. *et al.*, 2016. Zika Virus Associated with Microcephaly. *New England Journal of Medicine*, 374(10), pp.951–958.
- Mumtaz, N. *et al.*, 2016. Zika Virus: Where Is the Treatment? *Current Treatment Options in Infectious Diseases*, 8(3), pp.208–211.
- Noppakunmongkolchai, W. *et al.*, 2016. Inhibition of protein kinase C promotes dengue virus

- replication. *Virology Journal*, 13(35), pp.1–13.
- Nowakowski, T.J. *et al.*, 2016. Expression analysis highlights AXL as a candidate zika virus entry receptor in neural stem cells. *Cell Stem Cell*, 18(5), pp.591–596.
- Nugent, E.K. *et al.*, 2016. Zika virus: Epidemiology, pathogenesis, and human disease. *The American Journal of the Medical Sciences*, 353(5), pp.466–473.
- De Oliveira, A.S. *et al.*, 2014. NS3 and NS5 Proteins: Important Targets for Anti-Dengue Drug Design. *J. Braz. Chem. Soc*, 25(10), pp.1759–1769.
- Passi, D. *et al.*, 2017. Zika Virus Diseases – The New Face of an Ancient Enemy as Global Public Health Emergency (2016): Brief Review and Recent Updates. *International Journal of Preventive Medicine*, 8(6), pp.1–9.
- Petersen, L.R. *et al.*, 2016. Zika Virus. *The New England journal of medicine*, 374(16), pp.1552–63.
- Plourde, A.R. & Bloch, E.M., 2016. A Literature Review of Zika Virus. *Emerging Infectious Diseases*, 22(7), pp.1–15.
- Ramharack, P. & Soliman, M.E.S., 2016. Zika virus drug targets: a missing link in drug design and discovery – a route map to fill the gap. *RSC Advances*, 6(73), pp.68719–68731.
- Saiz, J. *et al.*, 2016. Zika Virus: the Latest Newcomer. *Frontiers in microbiology*, 7(496), pp.1–19.
- Sampath, A. & Padmanabhan, R., 2009. Molecular targets for flavivirus drug discovery. *Antiviral Research*, 81(1), pp.6–15.
- Screaton, G. *et al.*, 2015. New insights into the immunopathology and control of dengue virus infection. *Nature Reviews Immunology*, 15(12), pp.745–759.
- Shapshak, P. *et al.*, 2016. Global virology I-identifying and investigating viral diseases. In *Global Virology I-Identifying and Investigating Viral Diseases*. New York: Springer Science+Business Media, pp. 477–500.
- Singh, R.K. *et al.*, 2016. Zika Virus - Emergence, evolution, pathology, diagnosis and control: current global scenario and future perspectives - A comprehensive review. *The Veterinary quarterly*, 2176(July), pp.1–43.
- Sirohi, D. *et al.*, 2016. The 3.8 Å resolution cryo-EM structure of Zika virus. *Science*, 352(6284),

pp.467–70.

- Sironi, M. *et al.*, 2016. Nonstructural Proteins Are Preferential Positive Selection Targets in Zika Virus and Related Flaviviruses. *PLoS Neglected Tropical Diseases*, 10(9), pp.1–18.
- Stahla-Beek, H.J. *et al.*, 2012. Identification of a Novel Antiviral Inhibitor of the Flavivirus Guanylyltransferase Enzyme. *Journal of Virology*, 86(16), pp.8730–8739.
- T. Naga Ravikiran , T. Nagamounika, Y.R.P., 2016. ZIKA VIRUS -A LATEST VIRAL PANDEMIC. *Indo American Journal of Pharmaceutical Research*, 6(6), pp.5699–5706.
- Tang, H. *et al.*, 2016. Zika virus infects human cortical neural progenitors and attenuates their growth. *Cell Stem Cell*, 18(5), pp.587–590.
- Tian, H. *et al.*, 2016. Structural basis of Zika virus helicase in recognizing its substrates. *Protein and Cell*, 7(8), pp.562–570.
- Tian, H. *et al.*, 2016. The crystal structure of Zika virus helicase: basis for antiviral drug design. *Protein & Cell*, 7(6), pp.450–454.
- Wahid, B. *et al.*, 2016. Zika: As an emergent epidemic. *Asian Pacific Journal of Tropical Medicine*, 9(8), pp.723–729.
- Wells, M.F. *et al.*, 2016. Genetic Ablation of AXL Does Not Protect Human Neural Progenitor Cells and Cerebral Organoids from Zika Virus Infection. *Cell Stem Cell*, 19(6), pp.703–708.
- White, D.N.J., 1977. Molecular Mechanics Practice, *RSC advances*, 12(1), pp.225–233.
- White, M.K. *et al.*, 2016. Zika virus: An emergent neuropathological agent. *Annals of Neurology*, 80(4), pp.479–489.
- WHO, 2016a. RAPID RISK ASSESSMENT Zika virus disease epidemic : potential association with microcephaly and Guillain-Barré syndrome (1 st update) Main conclusions. *European Centre for Disease Prevention and Control*, pp.1–20.
- WHO, 2016b. *Zika virus, Microcephaly and Guillain-Barré syndrome*, Brazil.
- Zhang, R. *et al.*, 2016. A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. *Nature.*, pp.54–60.
- Zhang, S. *et al.*, 2016. Neutralization mechanism of a highly potent antibody against Zika virus. *Nature Communications*, 7, p.13679.

Zou, G. *et al.*, 2011. Functional analysis of two cavities in flavivirus NS5 polymerase. *Journal of Biological Chemistry*, 286(16), pp.14362–14372.

CHAPTER 3

3. Molecular Modeling and Computational Approaches to Biomolecular Structure and Drug Design

3.1 Introduction

Molecular modeling is one of the most rapid developing scientific fields, as it comprises of a wide range of theoretical and computational tools used to model and simulate small chemical and biological systems with the purpose of understanding their behavior at an atomistic level (Kore *et al.* 2012).

While experimental techniques can significantly demonstrate the mechanism of action of a biological system, the extensive labor, time reservations and financial shortfalls have led research communities toward enhanced computational alternatives (Cramer 2004; Lu *et al.* 2012). The discipline of computational chemistry forms part of the nucleus of molecular modeling, allowing for significant medical breakthroughs due to immense improvements in computer hardware and software over recent decades (Jensen 2007). Starting in the 1960s and progressing rapidly since the late 1980s, these computational techniques have provided a robust platform for biomolecular structure analysis and drug discovery (Leach 2001; Song *et al.* 2009).

Rational drug design is based on the fundamental knowledge that the activity of a drug is obtained from the binding of the compound to a molecular pocket of the biological target. The drug's chemical and geometric stability at the molecular pocket is complementary to successful activity. The computational methods used in rational drug design and structure analysis include: protein modeling (homology modeling), sequence diversity analysis, virtual screening and molecular docking (Kore *et al.* 2012; Huang *et al.* 2010).

There are two essential molecular modeling principles (Figure 3.1) that may be used to establish the energetics and conformational changes to the drug-target system:

- Quantum Mechanics and
- Molecular Mechanics.

By combining the above molecular modeling principles with molecular dynamic simulations, the target's flexibility and inhibitor binding landscape may be analyzed (Lewars 2003).

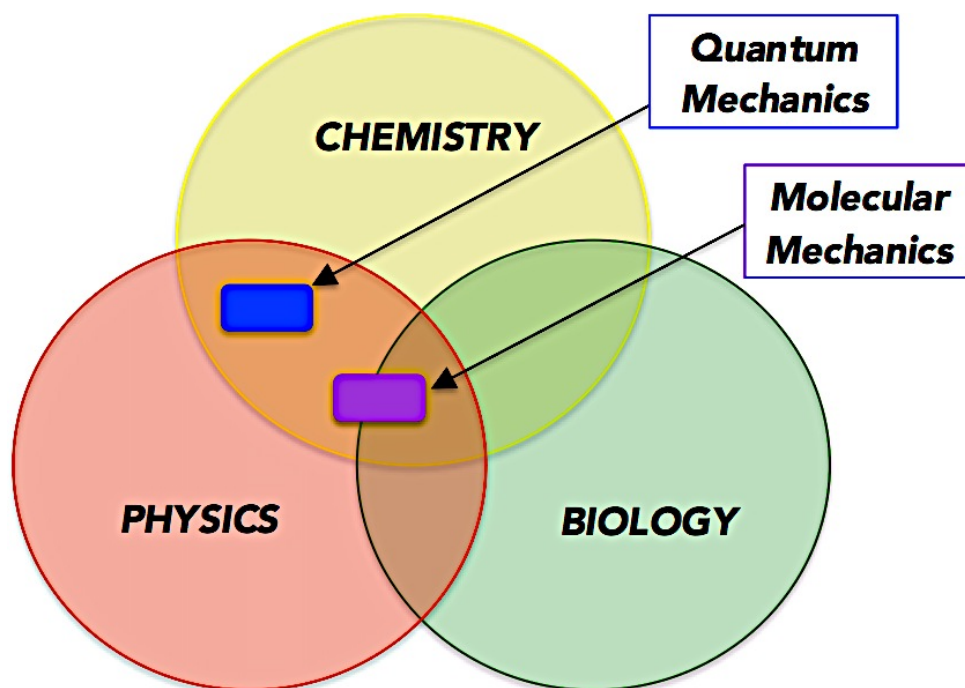


Figure 3.1: The scientific domains in which Applications of Quantum and Molecular Mechanics fit into (Prepared by Author).

In this chapter, quantum mechanics, molecular mechanics and molecular dynamic simulations will be elaborated on, thus providing insight into the rationale behind the chosen energy descriptors for this study. The principle behind each of the computational tools employed in the study will also be further explained.

3.2 The Principle of Quantum Mechanics

Quantum mechanics (QM) is one of the most successful branches of physics. The Principles of QM were developed during the early 20th century, where 2 types of QM were established. The initial development of Matrix mechanics was by German scientists: Planck, Born, Jordan and Heisenberg as well English-borne Dirac. Later, in 1926, Erwin Schrödinger developed wave mechanics, which now plays a fundamental role in the understanding of quantum phenomena (Trabesinger 2009).

The Quantum theory explains the behavioral characteristics of sub-atomic particles, such as electrons, at a nano-scopic level (Jensen 2007; Atkins & Friedman 2011). The phenomena of QM play important roles in biological processes of molecular biology such as bond forming/breaking, atomic transfer and electron excitation. Theoretically, QM calculations can predict any property of an individual system in a 3D- space. Electrons are mapped using the continuous electron density method and the energetics of the system is calculated using Schrödinger's wave function theory. For larger systems, electron density may be calculated using the earlier released Born-Oppenheimer approximation theory (Shen *et al.* 2016; Jakobsson 2001).

Provided below is the basic principle of Schrödinger's wave function and Born-Oppenheimer approximation theory:

3.2.1 The Schrödinger Wave Function

In January 1926, Austrian physicist, Erwin Schrödinger proposed the quantum mechanical model of the atom. Expanding on the Bohr atom model, which proposes that electrons are arranged in concentric circular orbits around a nucleus, Schrödinger utilized mathematical equations to describe the probability of locating an electron on an exact path. The model is portrayed as a nucleus that is surrounded by an electron cloud of high and low densities. According to quantum mechanics, all particles are described as a wave function with no defined position or momentum until they are observed. The probability of each possible observation may be determined by the wave function (Leach 2001; Atkins & Friedman 2011).

The Schrödinger equation forms the fundamental core of QM, as Schrödinger himself found that by adding the properties of an atom, being the mass and charge, to the equation, he was able to predict a series of shapes showing the wave pattern of electrons in an atom (Bahrami *et al.* 2014).

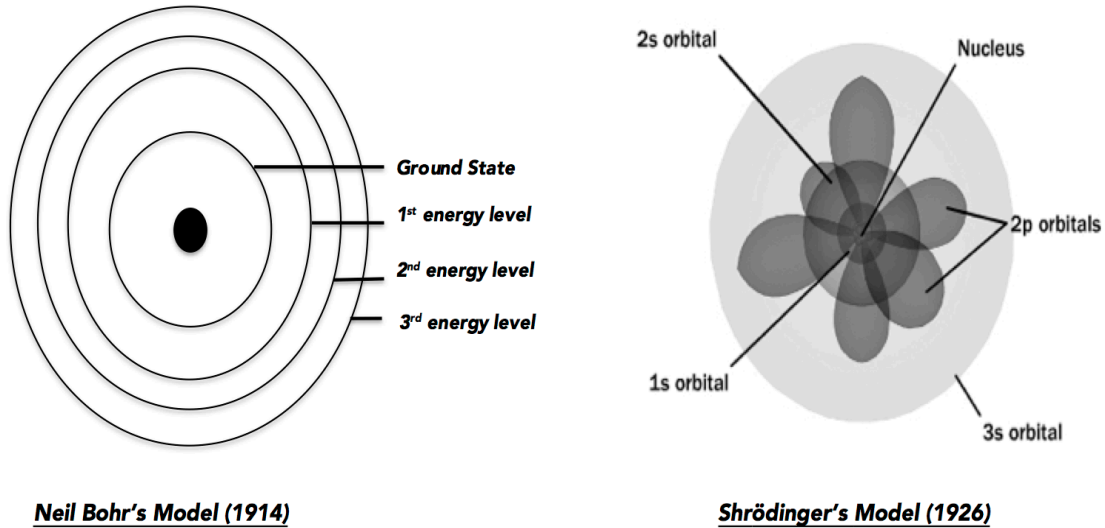


Figure 3.2: The Bohr Model demonstrated the atom to have a positively charged nucleus that was orbited by negatively charged electrons. This model was corrected by the equation, which evidenced electrons to have wave functions dependent on mass and charge of the atom. The two models are the fundamentals of what we now know as Quantum mechanics (Prepared by Author).

The Schrödinger wave equation:

$$\mathbf{H}\Psi = \mathbf{E}\Psi \quad (\text{Eq 3.2.1})$$

Where H is called the Hamiltonian operator (contains derivatives with respect to atom location), E is the energy eigenvalues of the system and ψ is the wave function. In order to replicate a relevant physical model of Schrödinger's equation, the wave function must be continuous, single valued, normalized and anti-symmetric. The molecular Hamiltonian operator is the sum of the atom's total potential energy (V) and kinetic energy (T):

$$\mathbf{H} = \mathbf{T} + \mathbf{V} \quad (\text{Eq 3.2.2})$$

Where H is defined as follows:

$$\mathbf{H} = \left[-\frac{\hbar^2}{8\pi^2} \sum_i \frac{1}{m_i} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] + \sum_i \sum_{<j} \left(\frac{e_i e_j}{r_{ij}} \right) \quad (\text{Eq 3.2.3})$$

The Schrödinger equation is highly complex, thus proving to be in-executable when solving for molecular systems, as it may contain thousands of atoms (Nakatsuji 2004; Barde *et al.* 2015; Bahrami *et al.* 2014). However, another QM theory, The Born-Oppenheimer Approximation, compensates for molecular rather than atomic structure.

3.2.2 The Born- Oppenheimer Approximation Theory

In 1927, physicists Max Born and J. Robert Oppenheimer proposed the Born-Oppenheimer approximation, which describes the uncoupling of the nuclei wave function to that of the electrons (Born & Oppenheimer 1927). Electrons are taken to be of lighter weight than that of nuclei, thus having increased velocity and move instantaneously to nuclei movement. Electron distribution within a molecule is therefore defined by the location of the nuclei (Liehr 1957; Ochkur 1965). This allows for the Schrödinger equation to be solved for the kinetic energy of the electrons alone, as the kinetic energy for the nuclei will remain constant.

The difference in velocities of the nuclei and electrons allow for the Born-Oppenheimer approximation to be applied, minimizing the complexity of the wave function of the Hamiltonian equation (Huang & Yi 2009). The simplified wave function:

$$\Psi(\mathbf{r}_{\text{elec}}) = \Psi(\mathbf{r}_{\text{elec}}) (\Psi(\mathbf{r}_{\text{nucl}})) \quad (\text{Eq 3.2.3})$$

Eq 3.2.1 is converted:

$$\mathbf{H}_{EN} \Psi(\mathbf{r}_{\text{elec}}) = \mathbf{E}_{EN} \Psi(\mathbf{r}_{\text{elec}}) \quad (\text{Eq 3.2.4})$$

Where H_{EN} denotes a difference between terms based activity to fixed nuclear positions (V_{NN}) or their activity to the non-fixed electron positions. Eq. 3.2.5 shows E_{EN} , which is derived from 2 sources being the fluctuating electron co-ordinates and fixed nuclear co-ordinates.

$$(H_{\text{el}} + V_{\text{NN}}) \Psi(\mathbf{r}_{\text{el}}) = E_{\text{EN}} \Psi(\mathbf{r}_{\text{el}}) \quad (\text{Eq 3.2.5})$$

We use the electronic Schrödinger equation to describe electronic motion within a molecule. The Approximation is seen to be more accurate when applied to ground electronic states. Once the equation has been solved, fixed positions of interest of the equilibrated conformation may be assessed and the potential energy surface and curve may be constructed (Matsika 2010; Woolley 1991; Jecko 2014; Lewars 2003).

3.2.3 Potential Energy Surface as an Application of Quantum Mechanics

The potential energy surface is an effective mathematical/graphical representation between molecular vibrational motions of a molecule, its geometry as well as its nuclear probability distribution by solving the time-dependent Schrödinger equation. The concept of potential energy surface arises from the Born- Oppenheimer approximation as explained above, whereby electrons vary according to the positional states of the nuclei so that the potential energy surface is taken as the potential of an atoms motion to collide with each other in a molecule (Atkins & Friedman 2011; Woolley 1991; Lewars 2003; Levitt *et al.* 1995). A potential energy surface displays high potential energy regions, indicating high-energy nuclear arrangements or molecular conformations and low energy regions indicating low nuclear energy conformations (Figure 3.3). This may be utilized in computational chemistry to identify the lowest energy state and the positional geometry of a molecule at this state (Jakobsson 2001; Jensen 2007).

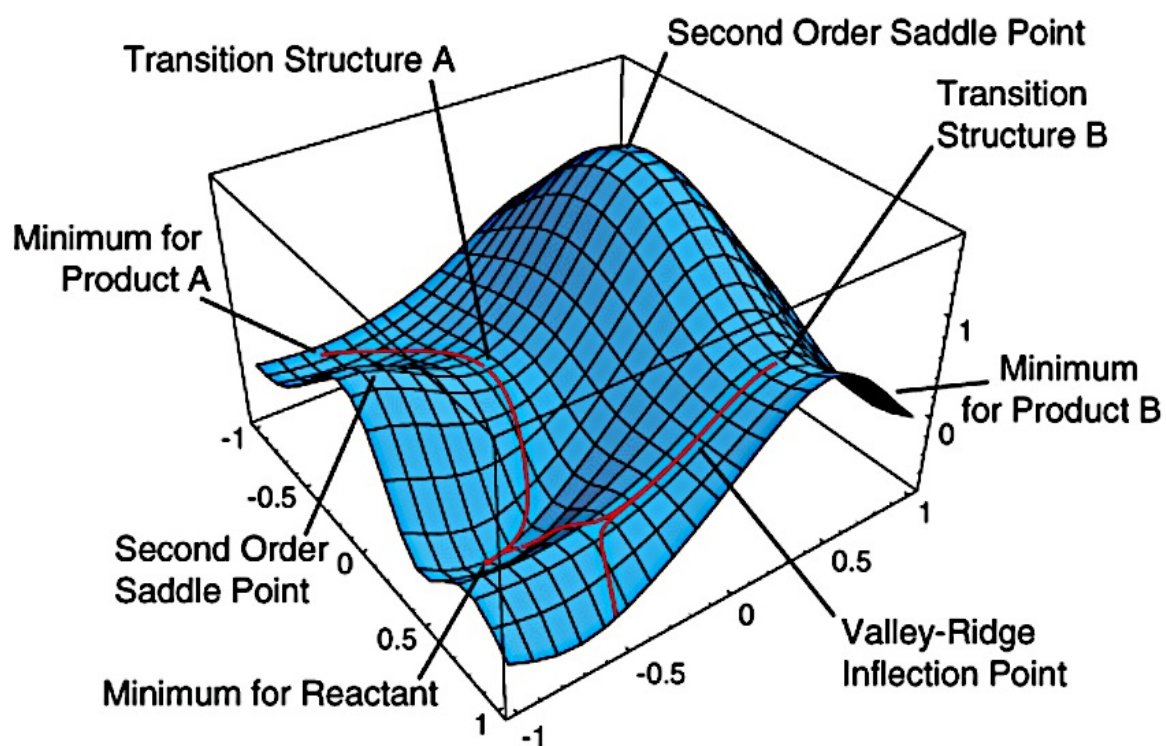


Figure 3.3: Graphical representation of a two-dimensional potential energy surface (PES) (University of California n.d.).

3.3 *The Principle of Molecular Mechanics*

One of the main difficulties in chemistry is to be able to understand the chemical characteristics of a compound, that is, its reactivity, solubility and stability. To measure these features, quantifiable dynamics need to be observed and analyzed from predicted molecular models (Boeyens & Comba 2001).

Molecular Mechanics (MM) may be defined as a set of models that utilize an empirical, algebraic, atomistic energy function for chemical systems. Also known as molecular force field methods, MM have become successful when dealing with large molecules that require multiple molecular dynamic calculations such as in biochemistry (Maseras & Morokuma 1995). It utilizes classical Newtonian mechanics to describe a large variety of molecular systems, from low molecular weight systems such as hydrocarbons, to large biomolecular complexes consisting of thousands of atoms such as proteins or membrane fragments (Vanommeslaeghe *et al.* 2014).

In Molecular mechanics, simple algebraic terms are used to express the total energy of a compound without needing to compute wave function or electron density as with quantum mechanics (Tsai 2002). Numerous techniques are utilized in rational drug design that identifies potentially desirable compounds prior to experimental testing. Molecular Mechanic simulations also allow for the construction of atomistic models based on favorable energy calculations (Poltev 2015).

3.3.1 *Potential Energy Function*

As mentioned above, atoms are classified as the “building blocks” in force field methods and electrons are not considered to be individual particles. This means that rather than solving the Schrödinger equation, explicit bonding information must be provided. In force field methods, molecules are described by a “ball and spring” model, with atoms of different sizes and bonds of different lengths. It was observed that different molecules might have structural similarity due to the atoms they are made up of. The concept was coined “atom types” and is dependent on the atomic number and chemical bonding holding it in place (Jensen 2007).

The potential energy function (PEF)/ force field energy of a molecular system may be classified in terms of a set of force field energy equations that are fundamentally based on classical Newtonian physics. These equations are able to calculate not only the energy of a system, but the “atom types” that make up the molecule as well (Jensen 2007; Tsai 2002).

The total potential energy comprises of the extended sum of all individual potential intra/inter molecular components, including:

1. Bond stretching (between directly bonded atoms)

$$\mathbf{E_r} = \sum \mathbf{K_r}(\mathbf{r} - \mathbf{r_0})^2 \quad (\text{Eq 3.3.1.1})$$

2. Angle bending (atoms bounded to same central atoms)

$$\mathbf{E_\theta} = \sum \mathbf{K_\theta}(\boldsymbol{\theta} - \boldsymbol{\theta_0})^2 \quad (\text{Eq 3.3.1.2})$$

3. Bond torsion

$$\mathbf{E_\phi} = \sum \mathbf{K_\phi}[1 + \cos(\mathbf{n\phi} - \boldsymbol{\phi_0})] \quad (\text{Eq 3.3.1.3})$$

4. Non-bonded interactions (van der Waals and electrostatic)

$$\mathbf{E_{nb}} = \left[\sum \sum \left(\frac{\mathbf{A_{ij}}}{\mathbf{r_{ij}^{12}}} - \frac{\mathbf{B_{ij}}}{\mathbf{r_{ij}^6}} \right) \right] + \left[\sum \sum \left(\frac{\mathbf{q_i q_j}}{\mathbf{D r_{ij}}} \right) \right] \quad (\text{Eq 3.3.1.4})$$

Where: K_r , K_θ , K_ϕ are force constants for bond, angle, and dihedral angle and r_0 , θ_0 , ϕ_0 are the equilibrium distance, angle and phase angle. Parameter r_{ij} is distance, while A_{ij} and B_{ij} are van der Waal parameters. D is the molecular dielectric constant; q_i and q_j are charge points.

In molecular mechanics, atoms are typically treated as spheres and bonds as springs. It is important to note that the properties mentioned above are easiest to describe mathematically when atoms are taken as spheres with characteristics radii. The final potential energy function equation is therefore:

$$\mathbf{E_{total}} = \mathbf{E_r} + \mathbf{E_\theta} + \mathbf{E_\phi} + \mathbf{E_{nb}} \quad (\text{Eq. 3.3.2})$$

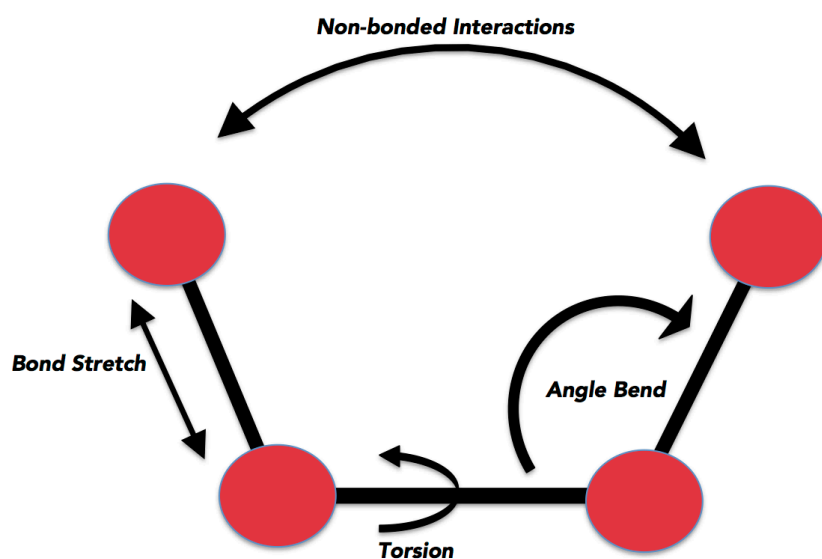


Figure 3.4: Diagrammatic representation of the total potential energy function of a molecule, as mentioned above (Prepared by Author).

There are many different force fields currently in use; however, they differ by the functional form of each energy term, the number of cross terms and the type of information used for fitting the parameters. Two general trends may be noted when designing the force fields:

1. Force fields used on large systems such as DNA or protein, have a relatively simple functional forms with no cross terms and use the Lennard-Jones potential as van der Waals energy. These are called harmonic/diagonal force fields.
2. Force fields used on small to medium size molecules have to maintain a high degree of accuracy. These have a number of cross terms and an exponential-type potential for van der Waals energy. These are called “Class II” force fields.

Examples of the most widely used and popular force fields are AMBER (Wang *et al.* 2004), GROMOS (Hermans *et al.* 1984), CHARMM (Brooks *et al.* 1983), OPLS-AA (Jorgensen *et al.*, 1996) and ENCAD (Levitt *et al.* 1995) (Monticelli *et al.* 2013). For the purpose of this study, the harmonic AMBER force field was utilized for the characterization of the molecular systems.

3.4 The Principle of Molecular Dynamics

Molecular dynamics (MD) first emerged in the late 1950s where Alder, Wainwright and Rahman developed simulation methods for the dynamics of liquids. The computational chemistry domain has progressed since then and from the early 1970s, molecular dynamics has become the most widely used method to study structure and dynamics of macromolecules such as DNA or protein (Tsai 2002). There are two types of simulation techniques, classical MD and Monte Carlo (MC). In recent years, numerous hybrid techniques have also been released. One of the major advantages of MD over MC is its ability to allow for dynamical properties of a system such as rheological properties and time-dependent responses (Nair & Miners 2014). Molecular dynamics is especially valuable in biochemistry and molecular biology as it affords the opportunity to identify and categorize, on an atomic scale, the dynamic events that may impact a biological properties of a system (Jarosaw Meller 2001).

Classical molecular dynamics incorporates Newton's equations of motion into its computational algorithm. Based on this highly evolved mathematical and physical algorithm, MD simulations provide high probability real-time conformational and mechanistic observations of many chemical reactions on an atomistic level (González 2011). An MD simulation allows us to study interacting particles of a system throughout a desired time-period, by producing a dynamical trajectory that may be then analyzed. The overall purpose of this computational technique is to utilize Newton's equations to solve and understand the energies and structural dynamics of a molecular network system. The following initial particle conditions are required:

1. Positions and velocities of each particle
2. A good force field to characterize the forces between atoms, e.g. AMBER or CHARMM
3. Boundary conditions that need to be engaged

The classical equation of motion may then be solved:

$$\mathbf{F}_i = m_i \frac{d^2 \mathbf{r}_i(t)}{dt^2} \quad (\text{Eq 3.4})$$

Where $\mathbf{r}_i(t)$ is the particle position vector, t is time-evolution, m is the mass of the particle and \mathbf{F}_i depicts the interacting force on the particle.

Molecular dynamics, according to Jakobsson *et al* (2001), may be rationalized into four continuous technical steps that are repeated millions of times to generate a trajectory (Jakobsson 2001). The steps are as follows:

1. The fundamental requirements (states) of the biomolecular system are defined:
 - The co-ordinates of each atom
 - The bond characteristics between each atom
 - The accelerations of atoms
2. Each atom's potential energy is computed.
3. The energies from step 2 are then utilized to solve the equations of motion.
4. The new "state" of the system needs to be saved and the atoms co-ordinates changed and step forward in the simulation is taken. The cycle then starts back at step 1.

Once the trajectory is fully generated, quantitative analysis of the system's time- evolution can proceed.

3.4.1 *Molecular Dynamics Post-Analysis*

Molecular dynamic trajectories are created from the production run of the simulation. The trajectories can be defined as sequential snapshots that are characterized by both positional co-ordinates and velocity vectors and detail the time evolution of the system in phase space (Likhachev *et al.* 2016; Jarosaw Meller 2001).

When choosing analytical software, three requirements are essential:

1. Qualitative visualization software that will not only display the trajectory's video clips, but also generates high quality snapshots/images.
2. The software should have prompt processors that will accommodate large volumes of data.
3. A variety of analysis options should be available on the one program.

The selected post-dynamic techniques and calculations should be dependent on the nature of the MD study; however, critical quantitative evaluation is necessary to support any visual systemization.

For the purpose of this study, the post dynamic analysis of the trajectories is critical to determining the:

1. energetic and conformational stability of the biomolecular system.
2. The characteristics of the system's small molecule binding landscape and the thermodynamic energy fluctuations along the system's clustered trajectory.
3. dynamic conformational features or variability of the biomolecular system.

3.4.1.1 System Stability

Convergence:

Convergence may be used to describe protein dynamics based on bond types and bond angle vibrations during the unfolding of a protein. This merging toward equilibrium and the representation of a final energetic and conformational plateau is essential for a MD trajectory to be accurate and reproducible (Amadei *et al.* 1999). It is at this plateau that the protein-ligand system is shown to display energetically stable conformations.

Root Mean Square Deviation (RMSD):

The deviation of a complex may be measured by the spatial difference between two static structures of the same trajectory. The RMSD of a trajectory is defined as:

$$\text{RMSD} = \left(\frac{\sum_N (\mathbf{R}_i - \mathbf{R}_1^0)^2}{N} \right)^{\frac{1}{2}} \quad (\text{Eq 3.5})$$

Where: N is the total number of atoms in the complex, \mathbf{R}_i is the vector position of the $\text{C}\alpha$ atom of particle i in the reference conformation which is computed after aligning the structure to the initial conformation (O) using the least square fitting.

The average RMSD may be calculated by taking the average over the number of frames in each trajectory and can be computed for the receptor, ligand and complex of a system (Kufareva & Abagyan 2012).

Radius of Gyration (RoG):

The radius of gyration in a protein may be defined as the root mean square distance of the atoms from their common centroid/center of gravity. This allows for the estimation of compactness of a protein complex along a trajectory. The RoG of a complex may be based on the following reaction:

$$r^2_{\text{gyr}} = \frac{(\sum_{i=1}^n w_i (r_i - r^-)^2)}{\sum_{i=1}^n w_i} \quad (\text{Eq 3.6})$$

Where: r_i is the position of the i th atom and r is the center weight of atom i .

The average RoG may be calculated by taking the average over the number of frames in a trajectory (Lobanov *et al.* 2008).

3.4.1.2 Thermodynamic Energy Calculations (Free Binding Energy)

Binding free energy calculations is an important end point method that may elucidate on the mechanism of binding between a ligand and enzyme, including both enthalpic and entropic contribution (Ylilauri & Pentikäinen 2013). Estimation of binding free energy leads to development of various algorithms and approaches including free energy perturbation, thermodynamic integration, linear interaction energy and molecular docking calculations, to mention a few.

Of all the free energy calculations, the Molecular Mechanics/Generalized Born Surface Area (MM/GBSA) and Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) methods have proven to be the most accurate and efficient in estimating binding free energies for biological macromolecules. Contradictory to molecular docking, both MM/GBSA and MM/PBSA do not rely on a large training set to define different parameters in each energy term. The above methods make use of a combination of molecular mechanics terms and the implicit solvent model to estimate the absolute free binding energy that is averaged over the number of frames in the trajectory (Genheden & Ryde 2015). The free binding energy (ΔG) computed by these methods for a protein system (complex, ligand and receptor) can be represented as:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (\text{Eq 3.7.1})$$

$$\Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS \quad (\text{Eq 3.7.2})$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (\text{Eq 3.7.3})$$

$$G_{\text{sol}} = G_{\text{GB/PB}} + G_{\text{SA}} \quad (\text{Eq 3.7.4})$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (\text{Eq 3.7.5})$$

Where: E_{gas} denotes the gas-phase energy, which consist of the internal energy E_{int} ; Coulomb energy E_{ele} and the van der Waals energies E_{vdw} . The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol} , was estimated from the energy contribution from the polar states, $G_{\text{GB/PB}}$ and non-polar states, G . The non-polar solvation energy, G_{SA} , was determined from the solvent accessible surface area (SASA), using a water probe radius of 1.4 Å, whereas the polar solvation, $G_{\text{GB/PB}}$, contribution was estimated by solving the GB/PB equation. S and T denote the total entropy of the solute and temperature respectively.

The MM/GBSA and MM/PBSA algorithms postulate quantifiable analysis of the binding affinity of the ligand to the protein and therefore are able to rationalize molecular docked structures (Godschalk *et al.* 2013).

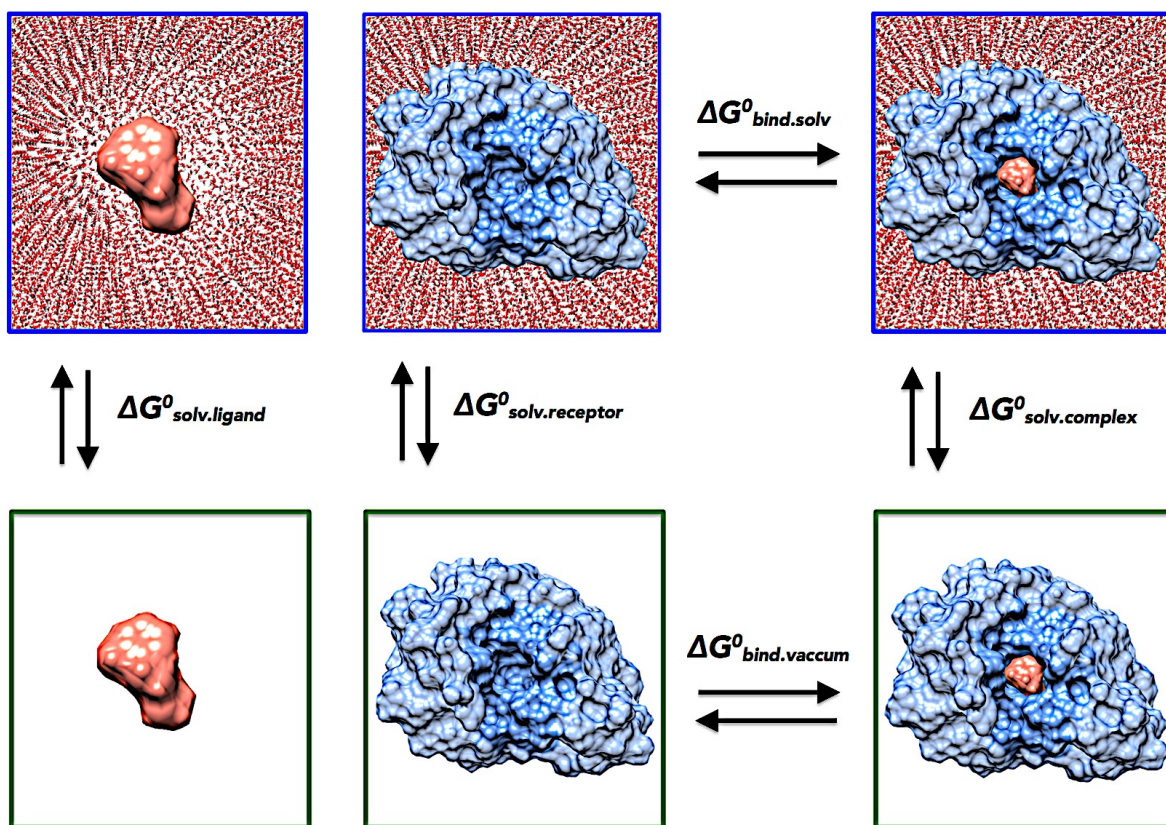


Figure 3.5: Diagrammatic representation of thermodynamic cycle as justified in the equations above (Prepared by Author).

3.4.1.3 Conformational Features of System

Root Mean Fluctuation (RMSF)

The root mean fluctuation (RMSF) of a protein measures residue's $C\alpha$ atom fluctuations based on the average protein structure along the system's trajectory. This extends to postulate the flexibility of regions of a protein based on the computed RMSF (Bornot *et al.* 2011). To calculate the standardized RMSF, the following equation is applied:

$$sRMSF = \frac{(RMSF_i - \overline{RMSF})}{\sigma(RMSF)} \quad (\text{Eq 3.8})$$

Where: $RMSF_i$ is the RMSF of the i^{th} residue, from which the average RMSF is subtracted. This is then divided by the RMSF's standard deviation to yield the resultant standardized RMSF.

The above method differs from RMSD and RoG as it is computed as the total residue fluctuation along the trajectory and is not analyzed at every frame in the trajectory.

Principal Component Analysis (PCA)

Principal component analysis (PCA) is defined as a covariance-matrix-based mathematical technique that is used to simplify the magnitude of the data generated from a MD simulation to understand correlated motions. In MD simulations of biomolecular systems, the PCA technique may be applied when measuring the atomic displacement and the loop dynamics of the protein.

The application of PCA in a MD simulation is known as “essential dynamics” as only fundamental motions of a data set are isolated from the millions of conformational snapshots. The conformational motions are then filtered from largest to smallest fluctuations and graphically depicted using a covariance matrix (Martinez & Kak 2001). The new set of defined co-ordinates are defined as the principal components of the data set and ordered such that the first 3-4 principal components have similar fluctuations as observed in the trajectory (David & Jacobs 2014). For the purpose of this study, the first 2 principal components were calculated and matrix covariance applied to evaluate the overall motion of the protein complexes.

Dynamic Cross Correlation (DCCM)

Dynamic cross correlation (DCCM) plots are used to quantify residue fluctuations either in or out of phase during a simulation. The cross correlation coefficient varies from -1 (completely anti-correlated motion) to +1 (completely correlated motion). The formula used to describe dynamic cross correlation is given below:

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{(\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle)^{\frac{1}{2}}} \quad (Eq\ 3.9)$$

The cross-correlation coefficient (C_{ij}) varies within a range of -1 to $+1$ of which the upper and lower limits correspond to a fully correlated and anti-correlated motion during the simulation

process. Where, i and j stands for i^{th} and j^{th} residue respectively and Δr_i or Δr_j represents displacement vectors correspond to i^{th} and j^{th} residue respectively.

Dynamic cross correlation maps have become very successful in quantifying residue motions that arise from ligand binding or in the occurrence of protein mutations (Kasahara *et al.* 2014; Tiberti *et al.* 2015).

3.5 Other Computer-Aided Drug Design Techniques Utilized in the Study

3.5.1 Homology Modeling

The initial step in molecular modeling and drug design is having a valid 3D structure, from X-ray crystallography, Nuclear Magnetic Resonance (NMR) or computational design using homology modeling (Soni & Madhusudhan 2017). The aim of homology modeling is to predict a three-dimensional (3D) model of a biological structure from a template sequence based on the structure of one or more homologous viral proteins of which crystal assemblies have been reported (Ramharack & Soliman 2016).

Homology modeling has played influential roles in many research areas and has aided in drug design by giving insights into spatial conformations and providing a structural template to construct novel drugs that are both specific and effective (Krieger *et al.* 2003).

In order to generate a 3D model of a biological target, a general procedure is followed, with validation at each step (Ramharack & Soliman 2016):

1. A target sequence needs to be identified and utilized to search (Blast) for homologous target sequences.
2. Template structure/s should be selected based on alignment length, sequence identity and structural identity.
3. Alignment between target and template sequences should be prepared.
4. Homology model should then be built using preferred computational software.
5. Validation of model may be verified using the predicted 3D structure and a Ramachandran plot.

The application of homology modeling may facilitate the design of low-resolution 3D structures that enable enhanced structural dynamic research of drug-target interactions on a molecular level (Hilbert *et al.* 1993).

3.5.2 Molecular Docking

One of the most popular tools utilized in computational drug design is molecular docking. The technique of molecular docking makes use of a multiple methods in the prediction of binding affinity and configuration of a complex. Ligand-receptor complexes exemplify the most general use of docking, although there are numerous studies that demonstrate protein-protein complex or drug delivery complexes such as nanoparticles or aptamers (Meng *et al.* 2011; Kroemer 2007).

There are two main steps involved in docking:

1. Sampling conformations of a ligand in the active site of protein- different algorithms may be used when sampling the numerous conformations of the docked complex: the “lock and key” model which describes the ligand and receptor as rigid structures, or the ligand may be flexible either through random or simulation-based methods. The latter algorithm is the most commonly used method as it allows for a more realistic fit of the ligand to the protein (Meng *et al.* 2011).
2. Ranking the different conformations by scoring function- the scoring function may be based on statistically preferred contacts, MM force fields or pre-existing protein-ligand binding affinities (Meng *et al.* 2011).

Over the past decade, there has been flood of molecular docking related publications and although these papers may add to the structural information about a biological target or new lead compound, there are still many inconsistencies that arise (Chen 2015). Frequent criticism associated with docking includes incorrect binding sites, choice of docked complex (conformational pose) and choice of small molecule (inhibitor or agonist) (Ferreira *et al.* 2015). Due to these concerns, all docked complexes in this study were verified with MD simulations where stability of the ligand at the active site was demonstrated.

3.5.3 Virtual Screening (VS)

Virtual screening (VS) plays a fundamental role in the drug discovery and development pipeline as the technique is defined by the assessment of extensive small molecule libraries in search of a new compound on the basis of a biological target. The VS techniques approach allows for the filtering of millions of small molecules to a manageable number of compounds that have the greatest chance as a lead drug. The method utilizes a wide variety of filters to identify biologically active alternatives to current inhibitors based on the “similar property principle”, which states that structurally similar molecules tend to have similar properties (Lionta *et al.* 2014; Vyas *et al.* 2008).

Virtual screening may be categorized into two approaches:

1. Structure-based virtual screening (SBVS) identifies energetically advantageous binding affinities of ligands into a target's active binding site. This allows for new insights on the nature of the active site and the protein-ligand interactions. The method identifies selective molecules from an extensive library of compounds to dock within a target's active site (Kumalo & Soliman 2016).
2. Ligand-based virtual screening generates libraries of compounds based on a known compound or compounds and its illustrative interactions with a particular target (Cele *et al.* 2016).

Of the approaches, SBVS has been shown to have similar inconsistencies as molecular docking and prove to be difficult when designing drugs for emerging diseases (such as this study). The LBVS generates large libraries of compounds and thus identifying accurate lead compounds is still challenging (Anderson 2003).

Pharmacophore based virtual screening (PBVS) has exhibited numerous benefits in computational hit identification and lead optimization. The approach uses pharmacophoric features based on a current inhibitor's functional groups (hydrogen bond donors, hydrogen bond acceptors, cations, aromatics, hydrophobic areas). These pharmacophoric features are then established as the criteria when searching through extensive small molecule libraries to identify a handful of compounds that may be validated as lead compounds. In this study, PBVS has been

employed as it has been evidenced to be more reliable than SBVS and LBVS (Kim *et al.* 2010; Sliwoski *et al.* 2014; Drie 2007).

Figure 3.6 summaries the computational tools carried utilized in this study.

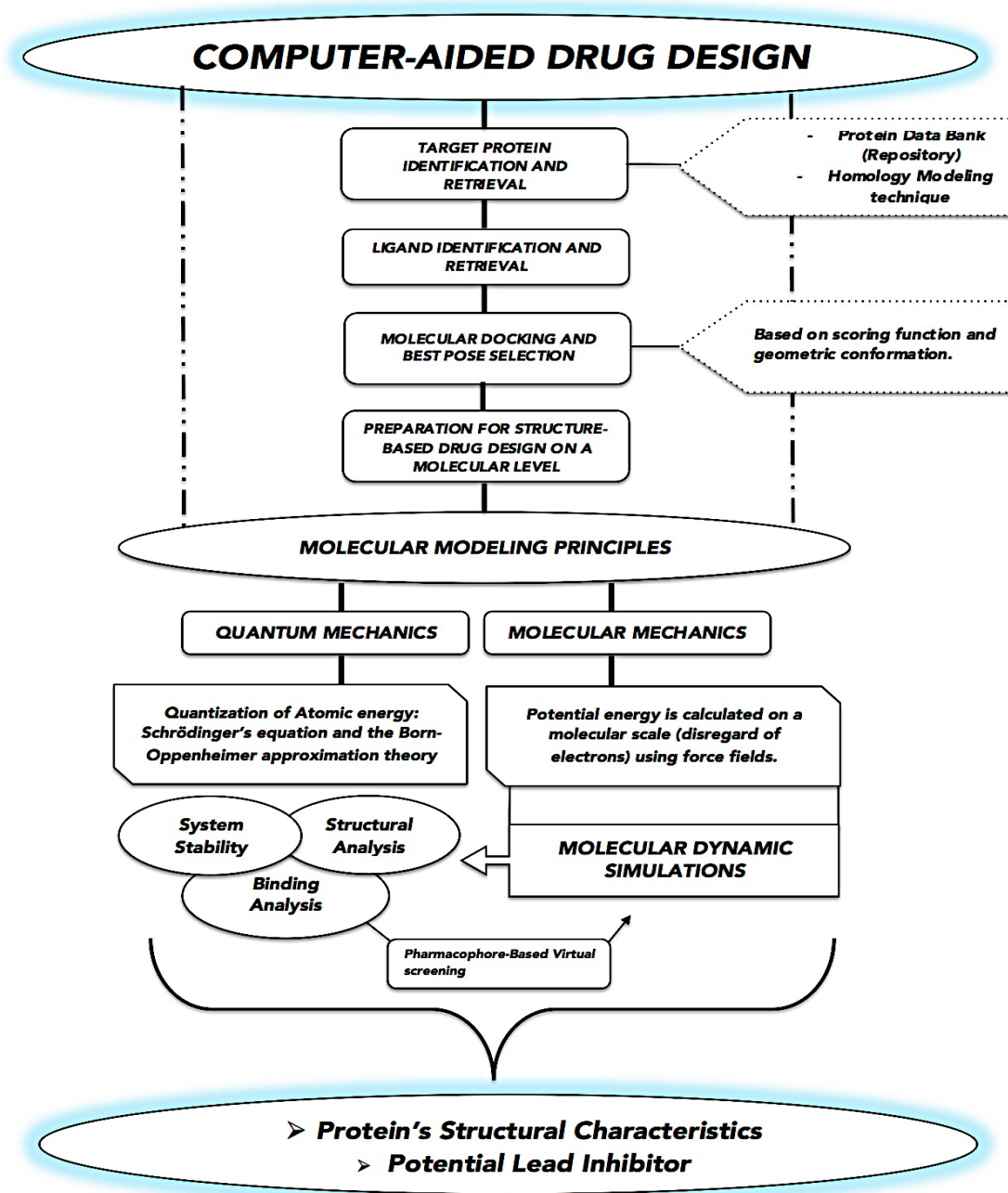


Figure 3.6: Summary of methods employed in this study (Prepared by Author).

References

- Amadei, A., Ceruso, M.A. & Di Nola, A., 1999. On the convergence of the conformational coordinates basis set obtained by the Essential Dynamics analysis of proteins' molecular dynamics simulations. *Proteins: Structure, Function and Genetics*, 36(4), pp.419–424.
- Anderson, A., 2003. The process of Structure-Based Drug Design. *Chemistry and Biology*, 10, pp.787–797.
- Atkins, P.W. & Friedman, R., 2011. *Molecular Quantum Mechanics*, 5(2), pp.654-659.
- Bahrami, M. *et al.*, 2014. The Schrodinger-Newton equation and its foundations. *New Journal of Physics*, 16, pp.1–17.
- Barde, N.P. *et al.*, 2015. Deriving time dependent Schrödinger equation from Wave-Mechanics, Schrödinger time independent equation, classical and Hamilton-Jacobi equations. *Leonardo Electronic Journal of Practices and Technologies*, 14(26), pp.31–48.
- Boeyens, J.C.A. & Comba, P., 2001. Molecular mechanics: theoretical basis, rules, scope and limits. *Coordination Chemistry Reviews*, 212(1), pp.3–10.
- Born, M. & Oppenheimer, J.R., 1927. Born-Oppenheimer approximation. *Annals of Physics*, 84, p.457.
- Bornot, A., Etchebest, C. & De Brevern, A.G., 2011. Predicting protein flexibility through the prediction of local structures. *Proteins: Structure, Function and Bioinformatics*, 79(3), pp.839–852.
- Brooks, B.R. *et al.*, 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, 4(2), pp.187–217.
- Cele, F.N., Muthusamy, R. & Soliman, M.E.S., 2016. Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: a study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. *Drug Design, Development and Therapy*, 10, pp.1365–1377.
- Chen, Y.C., 2015. Beware of docking! *Trends in Pharmacological Sciences*, 36(2), pp.78–95.
- Cramer, C.J., 2004. *Essentials of Computational Chemistry Theories and Models*, Springer, pp.1-547.

- David, C.C. & Jacobs, D.J., 2014. Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. *Methods Molecular Biology*, 1084, pp.193–226.
- Drie, J.H., 2007. Computer-aided drug design: The next 20 years. *Journal of Computer-Aided Molecular Design*, 21(10–11), pp.591–601.
- Ferreira, L.G. *et al.*, 2015. Molecular docking and structure-based drug design strategies, *Molecules*, 20(7), pp.13384-13421.
- Genheden, S. & Ryde, U., 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*, 10(5), pp.449–61.
- Godschalk, F. *et al.*, 2013. Comparison of MM/GBSA calculations based on explicit and implicit solvent simulations. *Physical chemistry chemical physics : PCCP*, 15(20), pp.7731–9.
- González, M.A., 2011. Force fields and molecular dynamics simulations. *Collection SFN*, 12, pp.169–200.
- Hermans, J. *et al.*, 1984. A consistent empirical potential for water–protein interactions. *Biopolymers*, 23(8), pp.1513–1518.
- Hilbert, M., Bohm, G. & Jaenicke, R., 1993. Structural relationships of homologous proteins as a fundamental principle in homology modeling. *Proteins: Structure, Function, and Bioinformatics*, 17(2), pp.138–151.
- Huang, H.J. *et al.*, 2010. Current developments of computer-aided drug design. *Journal of the Taiwan Institute of Chemical Engineers*, 41(6), pp.623–635.
- Huang, X.L. & Yi, X.X., 2009. Born-Oppenheimer approximation in open systems. *Physical Review A - Atomic, Molecular, and Optical Physics*, 80(3).
- Jakobsson, E., 2001. *Computational Biochemistry and Biophysics Edited by Oren M. Becker (Tel Aviv University), Alexander D. MacKerell, Jr. (University of Maryland), Benoît Roux (Cornell University), and Masa-katsu Watanabe (Wavefunction, Inc.). Marcel Dekker: New York and Ba*, *Journal of the American Chemical Society*, 123(50), pp.12745-12745.
- Jarosaw Meller, 2001. Molecular Dynamics. *Encyclopedia of Life Sciences*, pp.1–8.
- Jecko, T., 2014. On the mathematical treatment of the born-Oppenheimer approximation. *Journal of Mathematical Physics*, 55(5).
- Jensen, F., 2007. *Introduction to Computational Chemistry 2nd* edition, John Wiley and Sons,

England, pp.1-583.

- Kasahara, K., Fukuda, I. & Nakamura, H., 2014. A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer-DNA complex. *PLoS ONE*, 9(11).
- Kim, K., Kim, N.D. & Seong, B., 2010. Pharmacophore-based virtual screening : a review of recent applications. *Expert Opinion Drug Discovery*, 5(3), pp.205–222.
- Kore, P.P. *et al.*, 2012. Computer-Aided Drug Design: An Innovative Tool for Modeling. *Open Journal of Medicinal Chemistry*, 2(4), pp.139–148.
- Krieger, E., Nabuurs, S.B. & Vriend, G., 2003. Homology Modeling. *Structural Bioinformatics*, 857, pp.507–508.
- Kroemer, R.T., 2007. Structure-based drug design: docking and scoring. *Current protein & peptide science*, 8(4), pp.312–328.
- Kufareva, I. & Abagyan, R., 2012. Homology Modeling. *Methods Molecular Biology*, 857, pp.231–257.
- Kumalo, H.M. & Soliman, M.E., 2016. Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery: A Case Study on the Identification of Novel -Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents. *Cellular and Molecular Bioengineering*, 9(1), pp.175–189.
- Leach, A.R., 2001. *Molecular modelling: principles and applications* 2nd Edition, Pearson Education, England, pp. 1-485.
- Levitt, M. *et al.*, 1995. Potential energy function and parameters for simulations of the molecular dynamics of proteins and nucleic acids in solution. *Computer Physics Communications*, 91(1–3), pp.215–231.
- Lewars, E., 2003. *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics* 2nd Edition, Springer Science Business Media, Berlin, pp.1-584.
- Liehr, A.D., 1957. On the use of the Born-Oppenheimer approximation in molecular problems. *Annals of Physics*, 1(3), pp.221–232.
- Likhachev, I. V., Balabaev, N.K. & Galzitskaya, O. V., 2016. Available Instruments for

- Analyzing Molecular Dynamics Trajectories. *The Open Biochemistry Journal*, 10(1), pp.1–11.
- Lionta, E. *et al.*, 2014. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Current topics in medicinal chemistry*, 14(16), pp.1923–38.
- Lobanov, M.Y., Bogatyreva, N.S. & Galzitskaya, O. V., 2008. Radius of gyration as an indicator of protein structure compactness. *Molecular Biology*, 42(4), pp.623–628.
- Lu, J. *et al.*, 2012. Computational drug discovery. *Acta Pharmacologica Sinica*, 33, pp.1131–1140.
- Martinez, A.M. & Kak, A.C., 2001. PCA versus LDA. *Transactions on Pattern Analysis and Machine Intelligence*, 23(2), pp.228–233.
- Maseras, F. & Morokuma, K., 1995. IMOMM: A new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structures and transition states. *Journal of Computational Chemistry*, 16(9), pp.1170–1179.
- Matsika, S., 2010. The Born-Oppenheimer approximation. *The Journal of chemical physics*, 133(22), p.224103.
- Meng, X.-Y. *et al.*, 2011. Molecular docking: a powerful approach for structure-based drug discovery. *Current computer-aided drug design*, 7(2), pp.146–57.
- Monticelli, L., National, F. & Monticelli, L., 2013. Force fields for classical molecular dynamics. *Methods in molecular biology*, 924, pp.197–213.
- Nair, P.C. & Miners, J.O., 2014. Molecular dynamics simulations: from structure function relationships to drug discovery. *In Silico pharmacology*, 2(4), pp.1–4.
- Nakatsuji, H., 2004. Scaled Schrodinger equation and the exact wave function. *Physical Review Letters*, 93(3), pp.30403–1.
- Ochkur, V.I., 1965. The Born-Oppenheimer method in the theory of atomic collisions. *Soviet Physics JETP*, 18(2), p.503.
- Poltev, V., 2015. *Handbook of Computational Statistics*, Springer, Germany, pp.1-563.
- Ramharack, P. & Soliman, M.E.S., 2016. Zika virus drug targets: a missing link in drug design and discovery – a route map to fill the gap. *RSC Advances*, 6(73), pp.68719–68731.

- Shen, L., Wu, J. & Yang, W., 2016. Multiscale Quantum Mechanics/Molecular Mechanics Simulations with Neural Networks. *Journal of Chemical Theory and Computation*, 12(10), pp.4934–4946.
- Sliwoski, G. *et al.*, 2014. Computational methods in drug discovery. *Pharmacological reviews*, 66(1), pp.334–95.
- Song, C.M., Lim, S.J. & Tong, J.C., 2009. Recent advances in computer-aided drug design. *Briefings in Bioinformatics*, 10(5), pp.579–591.
- Soni, N. & Madhusudhan, M.S., 2017. Computational modeling of protein assemblies. *Current Opinion in Structural Biology*, 44(June), pp.179–189.
- Tiberti, M., Invernizzi, G. & Papaleo, E., 2015. (Dis)similarity Index to Compare Correlated Motions in Molecular Simulations. *Journal of Chemical Theory and Computation*, 11(9), pp.4404–4414.
- Trabesinger, A., 2009. History of quantum theory: The short version. *Nat Phys*, 5(6), p.383.
- Tsai, C.S., 2002. Molecular modeling: molecular mechanics. In *An introduction to computational biochemistry*. pp. 285–314.
- Vanommeslaeghe, K. *et al.*, 2014. Molecular mechanics. *Current pharmaceutical design*, 20(20), pp.3281–92..
- Vyas, V. *et al.*, 2008. Virtual screening: A fast tool for drug design. *Scientia Pharmaceutica*, 76(3), pp.333–360.
- Wang, J. *et al.*, 2004. Development and testing of a general Amber force field. *Journal of Computational Chemistry*, 25(9), pp.1157–1174.
- Woolley, R.G., 1991. Quantum chemistry beyond the Born-Oppenheimer approximation. *Journal of Molecular Structure: THEOCHEM*, 230(C), pp.17–46.
- Ylilauri, M. & Pentikäinen, O.T., 2013. MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. *Journal of Chemical Information and Modeling*, 53(10), pp.2626–2633.

CHAPTER 4

Zika Virus Drug Targets: A Missing Link In Drug Design And Discovery – A Route Map To Fill The Gap

Pritika Ramharack¹ and Mahmoud E. S. Soliman^{1*}

¹Molecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

* Corresponding author: Mahmoud E.S. Soliman, email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Abstract:

Zika Virus is an emerging virus that has been defined by the World Health Organization as a serious global biological-threat. Zika virus is an arbovirus from the *flavivirus* genus that is linked to microcephaly after prenatal transmission from the infected mother and most recently Gullian-Barrè Syndrome. The need for innovative research methods is urgent due to the ambiguity surrounding Zika virus. The lack of experimental data regarding potential drug targets, strategies for design and drug resistance has prompted us to provide a comprehensive framework with structured theoretical and technical guidelines on potential drug targets, modeling and design of inhibitors against the virus, thus assisting and encouraging scientists from different research domains to fill the gap in this research area. We have also represented a 3D homology model of the ideal Zika viral target, the non-structural protein 5, identified the active binding sites of each domain of the protein and found potential compounds that may act as inhibitors. This report will be immensely beneficial toward the design of Zika virus drug inhibitors.

Key words:

Zika virus drug targets, computer-aided drug design

1. Introduction:

Zika Virus (ZIKV) is a re-emerging arthropod-borne virus that is predominantly found in the tropics, however, rapidly evolving climate conditions coupled with increasing distribution of *Aedes* mosquito vectors and emerging modes of transmission of the virus have increased the potential to cause outbreaks in previously unaffected areas (1). The virus is a member of the Spondweni serocomplex of the genus *flavivirus*, family *flaviviridae*. Other arboviruses related to ZIKV include Dengue virus, Japanese encephalitis viruses and West Nile virus (2,3).

The first cases of the ZIKV infection were reported in Nigeria in the 1950's. Since then, ZIKV has shown erratic cases in countries such as Uganda, Tanzania, Egypt, Gabon, and in parts of Asia including India and Indonesia, with the most devastating pandemic occurring in Brazil in 2015 (4,5). Since the outbreak in Brazil, infection has spread rapidly throughout South America and Mexico, with Colombia being one of the most-affected countries with over 20,000 suspected cases (6). As of June 2016, thirteen countries have reported Central Nervous System (CNS) malformations such as microcephaly and Guillain-Barré syndrome (GBS) which may potentially be linked to ZIKV; during the recent circulation of the virus, eight countries had reported cases of GBS, where laboratory testing confirmed ZIKV infection in a number of those cases (7). Globally, the prevalence of ZIKV infection may be greatly underestimated (Figure 4.1) due to the recently verified prenatal and sexual transmission in humans (8), as well as the abstruseness surrounding the pathogenicity and thus, in turn the search of inhibitors of this “neglected disease”.

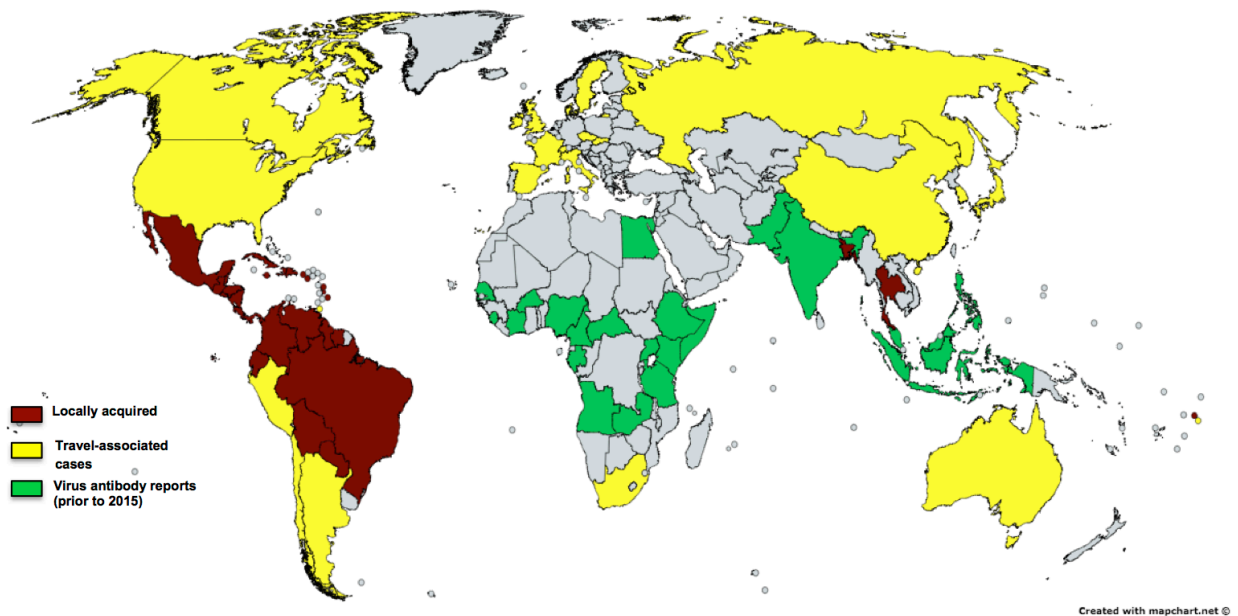


Figure 4.1: Global reports of ZIKV transmission, infection and sporadic viral antibody reports prior to 2015, as of April 2016 (Adapted from Centers for Disease Control and Prevention).

Most cases of ZIKV have reported febrile flu-like conditions that may be mistaken for other viral infections such as yellow fever. Other symptoms include swollen lymph nodes, maculopapular skin rashes and joint pains (5,9,10). Current research has raised concerns that the virus could cause dramatic increases in microcephaly in newborns after prenatal transmission (6,11–13). Complications associated with prenatal infection encompass fetal growth restriction, neurological and ocular abnormalities, intracranial calcification and in some cases perinatal death or stillbirth (10,14).

The virus is transmitted via an *Aedes* mosquito vector, congenital and perinatal transmission, as well as sexual intercourse (8,11,14–17). Studies have also reported transmission via blood transfusion and laboratory exposure (8,18). Commercial assays have been utilized in the diagnosis of ZIKV infection, including Real Time Polymerase Chain Reaction (RT-PCR) kits and IgM-based Enzyme-Linked Immunosorbent Assay (ELISA) (10,19). Sample DNA and RNA for

these kits may be extracted from blood serum, semen, amniotic fluid, plasma, saliva and urine. In Dengue infection, NS1 protein may be detected in a host's blood serum at the onset of clinical symptoms, this may prove to be another method by which ZIKV infection may be identified (16,20–22).

Although recent publications have described the global spread, pathogenicity and bioinformatics of ZIKV and its comparison between other flaviviruses including Dengue, West Nile, Yellow fever and Japanese Encephalitis virus (1,3,8,10,13,19,20,23–30), fundamental research into ZIKV small molecule drug design will be key in developing inhibitors of target proteins of the virus. Ekins *et al* (2016) described possible drug discovery and potential homology models of multiple proteins of ZIKV, however, despite the execution of research methods, there are currently no known FDA approved drugs of ZIKV (31). This prompted us to conduct a concise route map depicting the steps taken toward identifying potential inhibitors of drug targets with no 3D crystal structure and by following the guide to create a homology model of a non-structural protein of the virus, thus assisting scientists from different research domains. These *in silico* guidelines will be vastly beneficial in aiding and accelerating ZIKV experimental drug discovery.

2. Overview of ZIKV protein assembly

ZIKV is an enveloped virus comprising of an 11 kilobase, single-stranded positive sense RNA genome consisting of 10,794 nucleotides encoding 3,419 amino acids (25). The open reading frame (ORF) of the 5' and 3' untranslated region (UTR) encodes a polyprotein that is cleaved into three structural proteins being the capsid, precursor membrane, and envelope. Seven non-structural (NS) proteins are also found in this assembly, namely, NS1, NS2A, NS2B, NS3, NS4A,

2K, NS4B, and NS5 (largest viral protein) (24). These viral assembly proteins may act as crucial molecules in drug discovery.

3. Potential biological drug targets Against ZIKV

3.1 Viral Drug Targets

Hughes *et al* (2010) stated that the potential of a protein as a therapeutic target and its effectiveness in drug design is essential for determining the biological utility of the protein (32). ZIKV contains viral proteins that may act as targets in drug design (Table 1).

Table 4.1: Potential ZIKV target proteins

PROTEIN	NCBI REFERENCE SEQUENCE	PDB CODE	RESIDUE COUNT
Structural Proteins			
Capsid	YP_009227206.1	5IZ7/5IRE	122aa
Precursor Membrane	YP_009227197.1	5IZ7/5IRE	168aa
Envelope	YP_009227198.1	5JHM/5JHL	500aa
Nonstructural Proteins			
NS1	YP_009227199.1	5IY3	352aa
NS2A	YP_009227200.1	Not available	226aa
NS2B	YP_009227201.1	Not available	130aa
NS3	YP_009227202.1	5JMT	617aa
NS4A	YP_009227203.1	Not available	127aa
2K	YP_009227209.1	Not available	23aa
NS4B	YP_009227204.1	Not available	251aa
NS5	YP_009227205.1	Not available	903aa

*aa- amino acid

The structural proteins of ZIKV, being the capsid, precursor membrane and envelope form the viral particle (33). The envelope (E) protein is the key surface protein as it is able to mediate various aspects, including binding and membrane fusion of the viral replication cycle, making it a significant target in drug design (5).

The nonstructural proteins participate in the replication of the RNA genome, virion assembly and invasion of the innate immune system. Of the nonstructural proteins, NS5, NS3 and NS1 have shown enzyme activity in other viruses of the *flavivirus* genus, creating ideal targets in inhibitor development (33).

NS5 is a bifunctional enzyme with a methyltransferase domain at its N-terminal end and a RNA-dependent RNA polymerase (RdRp) at its C-terminal end. Both N- and C-terminal domains contain an *S*-Adenosyl-methionine-dependent MTase core structure that folds into an $\alpha/\beta/\alpha$ sheet cradled between the N- and C-terminal subdomains (34). The protein engages in virus-host interactions and actively interacts with the host environment (1). To our knowledge, there is currently no available 3D crystal structure of the ZIKV NS5 protein.

The NS3 protein is a multifunctional, viral replication protein. The protease comprises of the N-terminal third of NS3 and nucleotide triphosphatase, the RNA triphosphatase, and finally the helicase components. NS3 can be considered a serine protease and contains a classical catalytic triad (His-51, Asp-75, Ser-135) (1,35). Agnihotri *et al* (2012) reported an *in silico* study in which a homology model of the flavivirus NS3 protein was created using 22 species of the flavivirus genus. This study is a critical tool in the understanding the flavivirus NS3 protein and thus the impact of the protein as a ZIKV target (36). The 3D-crystal structure of the NS3 Helicase protein has recently been reported in *Protein and Cell* where a conserved triphosphate pocket and a positively-charged tunnel were identified to be critical for the hydrolysis of nucleoside triphosphates and the accommodation of RNA respectively (37).

The 3D crystal structure of the noteworthy NS1 glycoprotein viral target was released earlier this year and was classified as a major antigenic marker of ZIKV infection (38). The NS1 is

synthesized as a monomer and dimerizes after post-translation modification in the replication cycle (39). The mature NS1 protein has significant immune evasive functions on the surface of cells, in the extracellular space and in cells by directly regulating the translation of viral RNA. Recent studies on Dengue virus have also evidenced NS1 to be associated with vascular leak and shock due to the disruption of TLR3 signaling pathways (40). Song *et al* (2016) reported NS1 to display a loop-surface interface with divergent electrostatic potential that may result in unique interactions with host machinery compared to that of other flaviviruses (38). This makes ZIKV NS1 an ideal target for chemoinformatics studies.

Inhibitors of these viral proteins may be designed using computer-aided drug design techniques to select structural molecules that may inhibit the replication of viruses such as ZIKV in a host.

3.2 ZIKV Host Targets

During the ZIKV replication cycle, host cell machinery is imperative in the translation of viral RNA and maturation of the replicated virus, thus targeting host proteins and pathways may be key to effective inhibition of viral replication.

One of the most researched host proteins in flavivirus infection is the endoplasmic reticulum glucosidase. These proteins allow for the cleavage of the terminal glucose from the glycan found at the glycosylation-site of the prM and envelope protein, thus leading to its maturation of the envelope protein (41). Studies have shown that many flaviviruses, including ZIKV, have a N-glycosylation at Asn154 (31,42). Castanospermine (CST) and deoxynojirimycin (DNJ) have

been established as potent inhibitors of alpha-glucosidases, thus preventing the early stages of glycosylation (43).

Hamel *et al* (2015) described the importance of dendritic cell-specific intracellular adhesion molecule 3- grabbing non-integrin (DC-SIGN), TIM and TAM receptors in the attachment and entry of ZIKV into the host cell before replication can occur (44). Small interfering RNA (siRNA) was also shown to completely inhibit the expression of the above proteins after 48 hours. Other informative publications on siRNA inhibition of flavivirus host machinery include a review by Hirsch (2010), an *in silico* based experimental study on Dengue virus by Noppakunmongkolchai *et al* (2016) and the silencing of the 3' UTR of ZIKV genome by Shawan *et al* (2015) (5,45,46).

A recent study published in June by *Nature* identified host endoplasmic reticulum-associated signal peptidase complex (SPCS) to be necessary for the proper cleavage of ZIKV prM and envelope proteins. The authors also demonstrated that the loss of SPCS signaling leads to a dramatic decrease in Dengue, Yellow fever, West Nile, JEV and Hepatitis C viruses. This study could be a critical cornerstone in targeting host proteins and pathways in ZIKV infection (47).

4. *In silico* studies conducted on ZIKV

Prior to 2016, only two *in silico* reports have been made toward the development of ZIKV inhibitors. Computational studies by Shawan *et al* (2014) showed the viral envelope glycoprotein to be the most immunogenic structural protein of the virus, thus, making it a candidate for vaccine development (5). Shawan *et al* (2015) also looked at small interfering RNA (siRNA) in gene

silencing of the 3' UTR of ZIKV genome (48). Following the Brazil outbreak, an influx of research output has flooded the scientific community. There has been numerous computational studies regarding ZIKV target proteins; crystals structures of the NS1, NS3, envelope and the 2 cryo-EM structures of the stable virus have been released (37,38,42,49,50). Ekins *et al* (2016) described *in silico* studies in both drug discovery and the homology models of both structural and nonstructural proteins (9,31). There have also been reports comparing the structural and sequence conformations of ZIKV to other flaviviruses including Dengue and West Nile viruses (51,52).

5. *In silico* route map toward the design and discovery of ZIKV inhibitors

Rational Drug Design may be classified into two groups, the first being the development of small molecules with the desired effects of the target, whose structural information is known and the second group being development of small molecules whose cell functions and structural information may not be known (53).

To date, there is no available 3D crystal structure of the ZIKV NS5 protein. This prompted us to create a route map (Figure 4.2) describing the techniques of the second group, thereby benefiting scientists from different research domains by informing them of fundamental computational techniques in the design of novel small drug molecules, allowing for increased output of validatory experimentation.

Computer-Aided Drug Design (CADD) represents computational methods and resources used in the design and discovery of new therapeutic solutions (54). Numerous bioinformatics tools and resources have been developed to advance the drug discovery process (55,56). The recent

improvements made in computational chemistry software, CADD and molecular dynamic simulations have led to innovative research methods in the pharmaceutical industry (57).

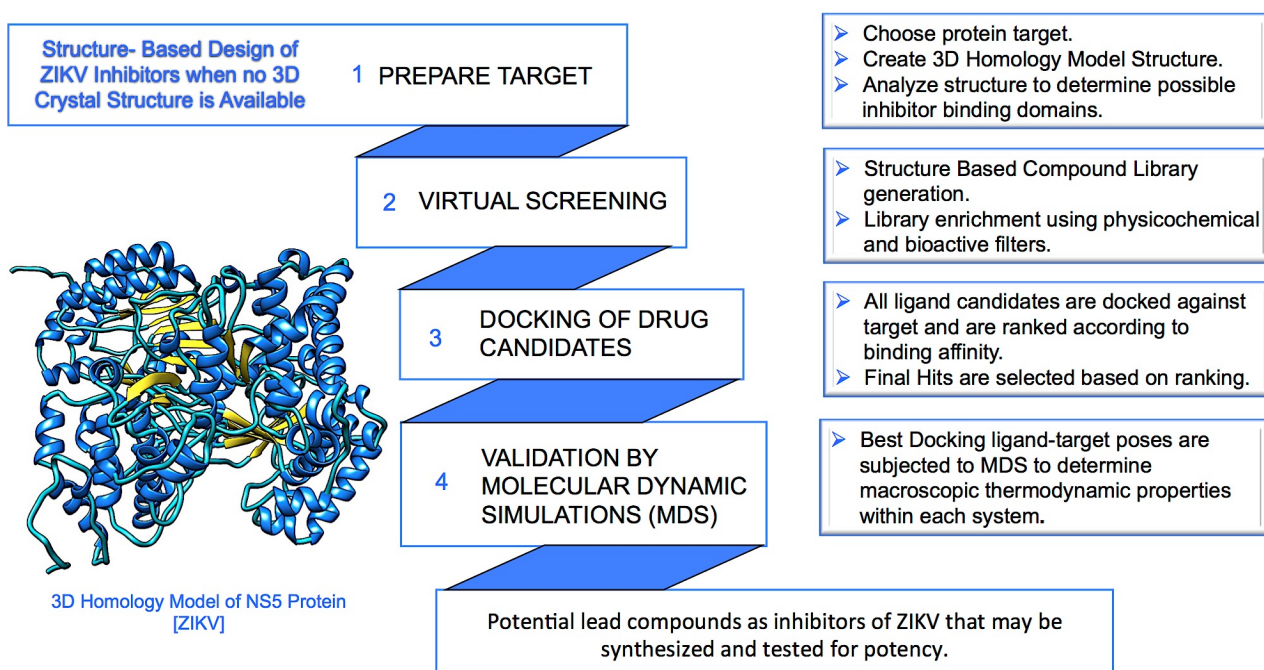


Figure 4.2: Route map toward the *in silico* design of ZIKV inhibitors using the homology modeled viral NS5 protein. Details on how the homology model was created are described under section 5.1.

The initial step of any modeling work is having a valid 3D structure, from X-ray crystallography, Nuclear Magnetic Resonance (NMR) or computational design using homology modeling. Homology modeling is used to predict and generate a plausible 3D structure of ZIKV's biological target from a template sequence based on the structure of one or more homologous viral proteins of which crystal assemblies have been reported (Figure 4.3) (58).

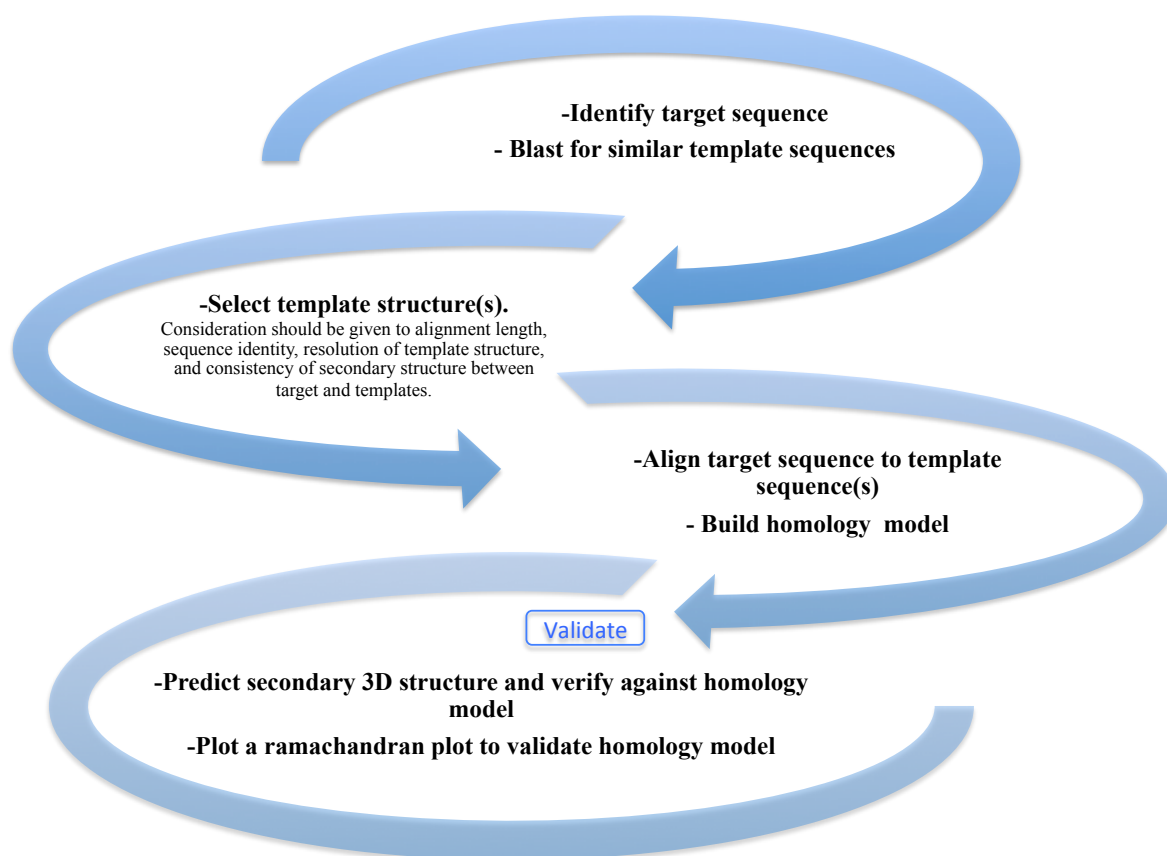


Figure 4.3: Protocol for building a homology model in our laboratory.

In order to create a ZIKV target 3D homology model, a typical procedure needs to be followed, with validation taking place at each step. Once the 3D structure has been generated and validated using 3D profiles and a Ramachandran plot (59), the predicted active binding site of the ZIKV target molecule may be identified. If the drug target is an enzyme, such as the NS3 or NS5 protein of the ZIKV viral assembly, designed chemical molecules may be able to fit within an active site pocket (56). The results establish the locality of possible binding pockets of the protein (60). After each pocket has been identified, we can identify the size of a pocket (volume, surface area and depth), possible interacting residues and surface atoms (61).

Subsequent to the ZIKV protein homology model and target site determination, several paths may be utilized in the development of inhibitors. Structure-based virtual screening will assist in searching through combinatorial chemistry libraries for molecules that may be potential inhibitors of the target protein and automatically dock these molecules into the 3D target's active pocket at a rapid rate (62). Thousands of molecules may be able to match the active site of the target protein, thus, a scoring function is utilized to rank ligands based on the free binding energy calculated after each docking pose (63,64). Molecules with the lowest free binding energy subsequent to screening may be used as inhibitor candidates, which may then be employed in a series of validity molecular dynamic simulations.

Molecular dynamic simulations calculate the trajectory of a generated docking pose by utilizing Newtonian mechanics (65). It is an important tool of CADD as it avoids analytic intractability in complex systems (57). Molecular dynamics is not essential in CADD but it can provide validation of docking results between a protein and its potential inhibitors (66).

By implementing *in silico* studies in the design of ZIKV protein inhibitors, putative drug-like compounds may be identified and their potency verified using *in vitro* and *in vivo* testing.

Studies report that *in vitro* testing of potential inhibitors may utilize cultured monkey cell lines such as LLC-MK2 and Vero (30). Delvichio *et al* (2016) also reported Chloroquine as potential ZIKV inhibitor in Vero, hBMEC, hNSC and mouse neurospheres. Dowall *et al* (2016) developed the first *in vivo* murine model, where adult female mice were subcutaneously inoculated with similar doses of ZIKV from natural infection of a mosquito bite (67,68). This model is a critical cornerstone in accelerated testing of new ZIKV inhibitors.

Larroca *et al* (2016) have made a significant contribution to the protection against the ZIKV virus by creating the first full-length prM-envelope DNA vaccine. The vaccine is currently undergoing clinical trials after the success in an *in vivo* study using infected mice. This vaccine may be the potential ‘holy-grail’ in ZIKV prevention (69).

5.1 A homology model for ZIKV NS5

In order for CADD of ZIKV to occur, a 3D crystal structure of a target protein is needed. Figure 4.4 shows the first account of a homology model for the ZIKV NS5 protein, which was created and validated as described in our previous publications (Figure 4.4) (60,70). The PDB coordinates of the homology model are provided as

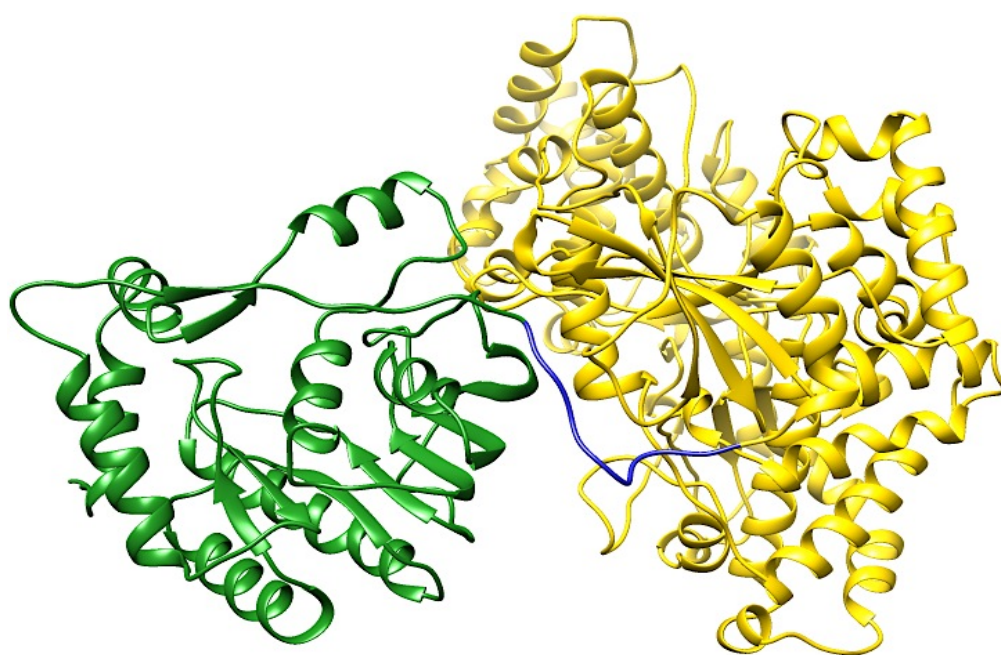


Figure 4.4: Homology model of ZIKV NS5 protein. The protein comprises of three domains, the N-terminal, methyltransferase domain (residues 1-262) (green), the inter-domain region (residues 263-272) (blue) and the C-terminal, RdRp domain (residues 273-903) (yellow).

5.2 Active site identification

Active site residues need to be identified for the docking of potential inhibitors to the active site pocket. The active site residues were determined using Chimera Multi-align Viewer and validated using the Site-Hound web program (71). Figure 4.5 highlights the best active sites and active site residue numbers of the NS5 protein (Figure 4.5).

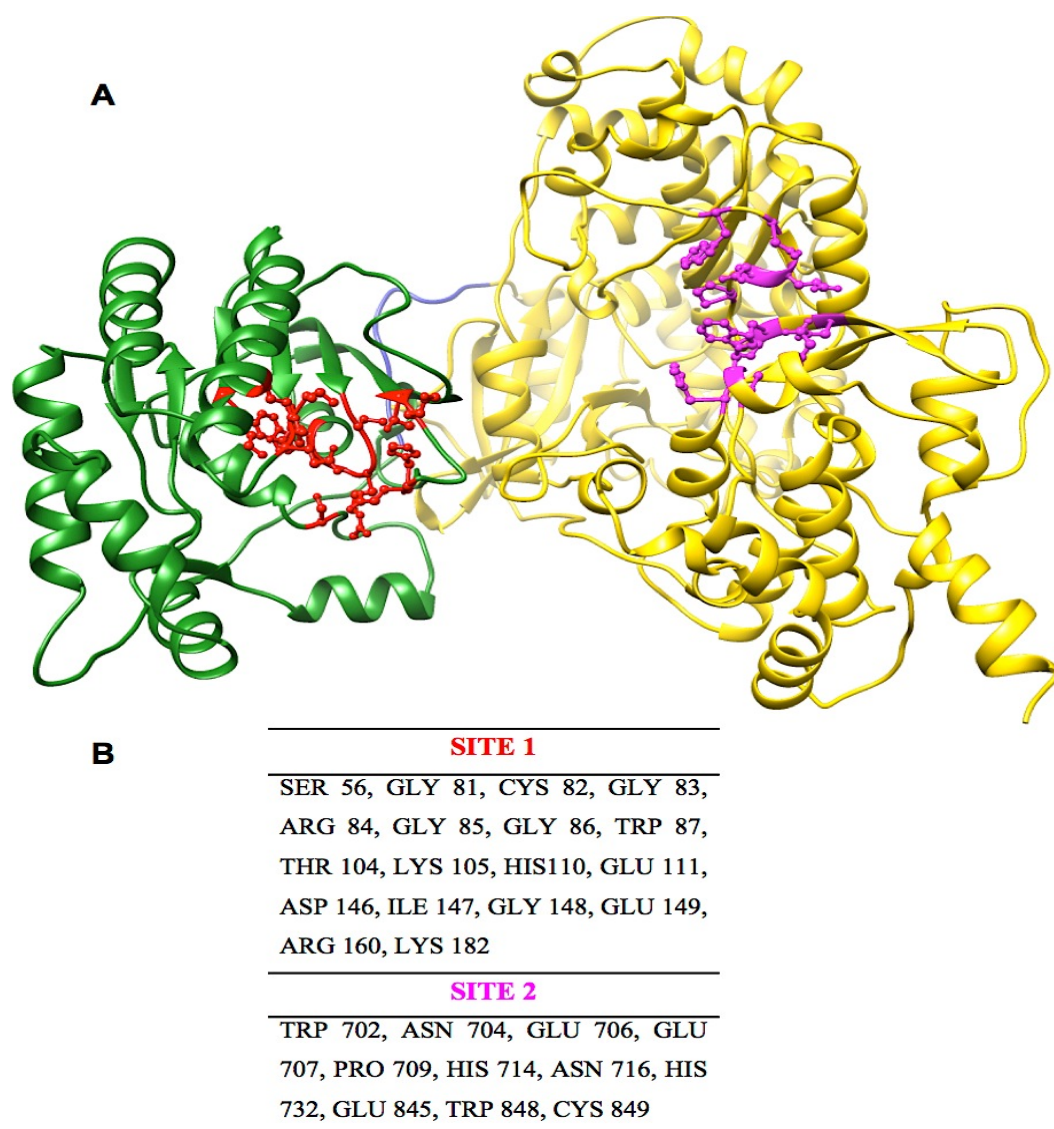


Figure 4.5: The potential binding sites, identified by Site-hound (71), of ZIKV NS5 protein. (A) Site 1 (Methyltransferase active binding site) (red) and Site 2 (RdRp active binding site) (magenta), (B) active binding site residues of the NS5 protein at Site 1 and Site 2.

This homology model will be implemented in the design of small molecules that may act as inhibitors of the NS5 protein, thus inhibiting the translation of viral RNA. Experimental drug therapy on other flaviviruses (41,72,73) may be used as a guide toward the identification of new specified small molecules that inhibit ZIKV replication.

5.3 Possible small molecule inhibitors of NS5 RdRp

Of the ZIKV target proteins, NS5 RdRp is one of the most favorable for drug discovery due to its role in viral replication (1). A study by Eyer *et al* (2016) looked at an *in vitro* study of nucleoside inhibitors against ZIKV and found one particular molecule, 2-C-Methyladenosine, to show promising inhibition of RdRp (74). The purine and hydroxymethyl structural features of 2-C-Methyladenosine were screened through ZINC database, criteria was imposed to ensure the inclusion of the maximum number of compounds, such that compounds had to have an xlog P between -4 and 5, a net charge 0, rotatable bonds between 0 and 8, a polar surface area of between 0 and 150, have hydrogen bond donors/acceptors between 0 and 10, and polar desolvation between 0 and 1 kcal/mol whereas compounds must have an apolar desolvation between -100 and 40 kcal/mol. Thereafter, the 4113 hits were downloaded and docked together with 2-C-Methyladenosine (Figure 4.6) at the RdRp active site and the nine best docked poses comprising the highest binding affinities were reported in Table 2. Table 3 shows the physical representation of the compounds. These compounds may be a basis for further validation and experimental verification.

Table 4.2: Representation of top ten compounds docking to NS5 RdRp

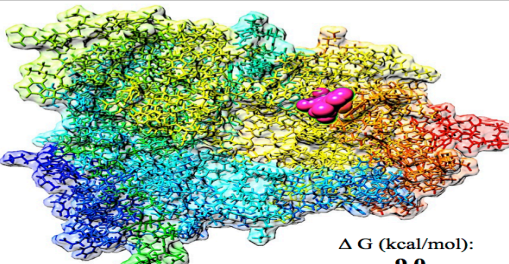
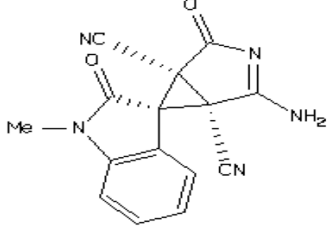
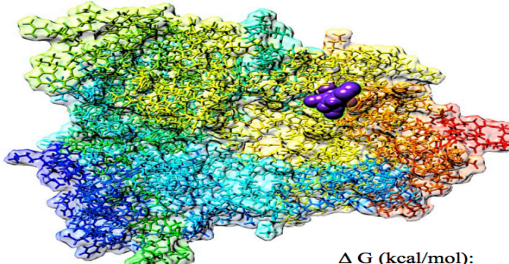
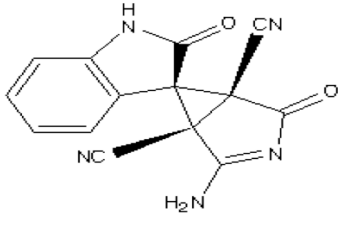
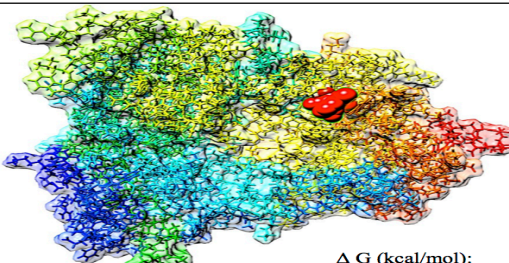
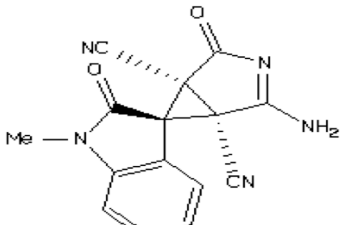
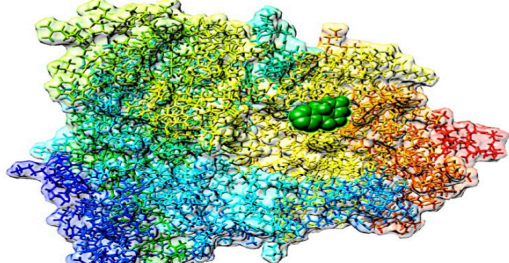
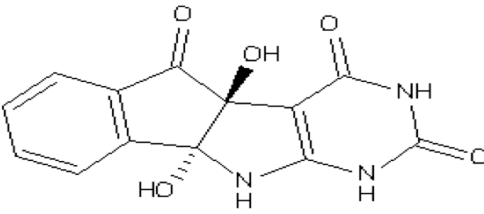
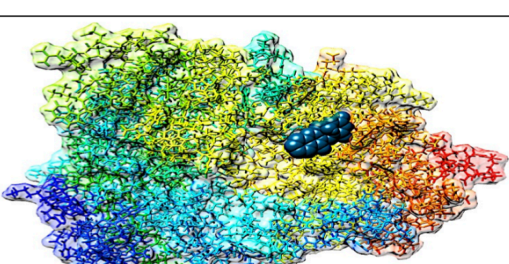
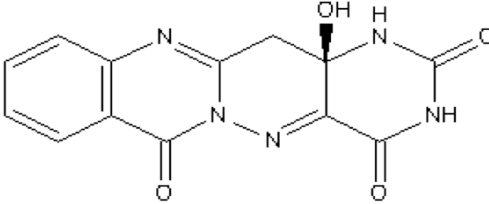
 <p>ΔG (kcal/mol): -9.0</p>	 <p>ZINC35325268</p>
 <p>ΔG (kcal/mol): -8.6</p>	 <p>ZINC14987423</p>
 <p>ΔG (kcal/mol): -8.3</p>	 <p>ZINC35325268</p>
 <p>ΔG (kcal/mol): -8.1</p>	 <p>ZINC00351019</p>
 <p>ΔG (kcal/mol): -8.1</p>	 <p>ZINC13633807</p>

Table 2 continued...

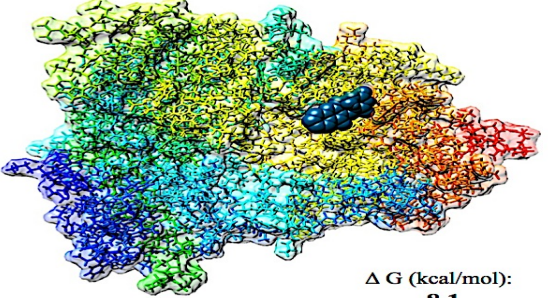
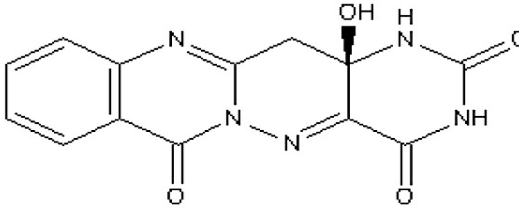
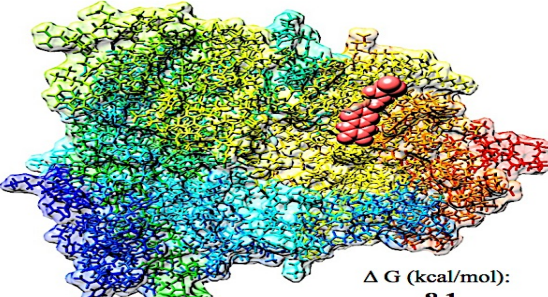
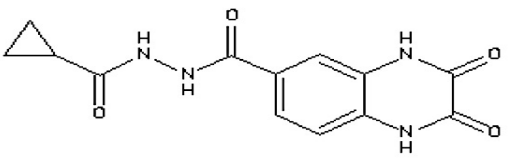
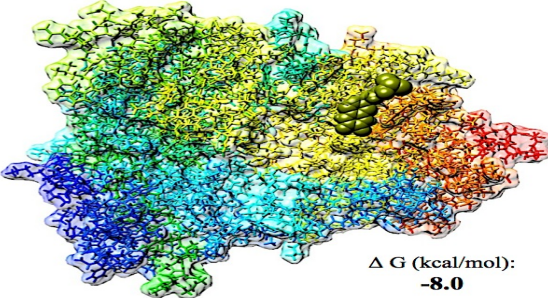
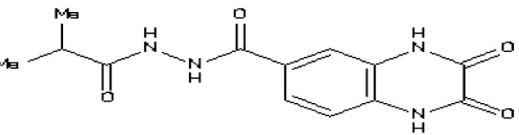
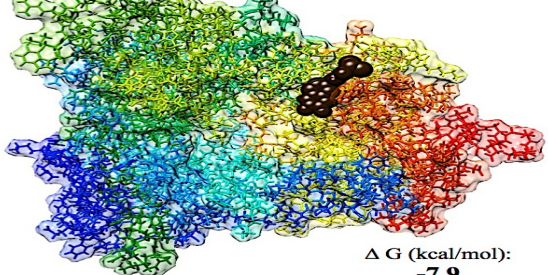
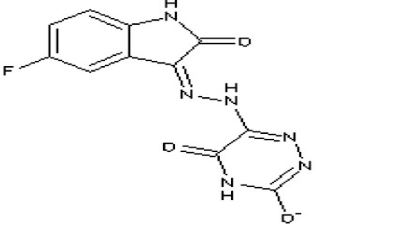
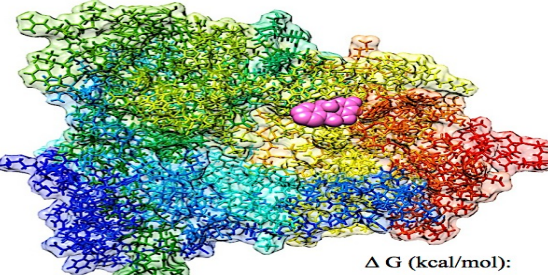
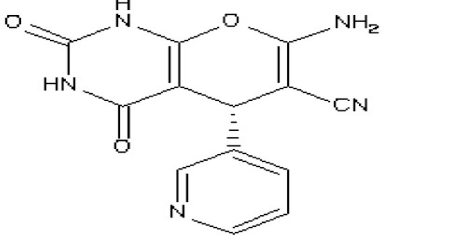
 <p>ΔG (kcal/mol): -8.1</p>	 <p>ZINC13633807</p>
 <p>ΔG (kcal/mol): -8.1</p>	 <p>ZINC40563785</p>
 <p>ΔG (kcal/mol): -8.0</p>	 <p>ZINC40563886</p>
 <p>ΔG (kcal/mol): -7.9</p>	 <p>ZINC13121997</p>
 <p>ΔG (kcal/mol): -7.9</p>	 <p>ZINC00043707</p>

Table 4.3: Physical Representation of top ten compounds displaying Molecular weight, xlogP, H-Bond Donors/Acceptors and Rotatable Bonds

ZINC ID	Molecular Weight	xlogP	Rotatable Bonds	H-Bond Donors	H-Bond Acceptors
ZINC35325271	291.27	-0.70	0	2	7
ZINC14987423	277.243	0.95	0	3	7
ZINC35325268	291.27	2.67	0	2	7
ZINC00351019	287.231	-5.47	0	5	8
ZINC13633807	299.246	-0.86	0	3	9
ZINC40563785	288.263	0.83	3	4	8
ZINC40563886	290.279	-0.89	3	4	8
ZINC13121997	289.206	-0.40	2	3	9
ZINC00043707	283.247	-1.01	1	4	8

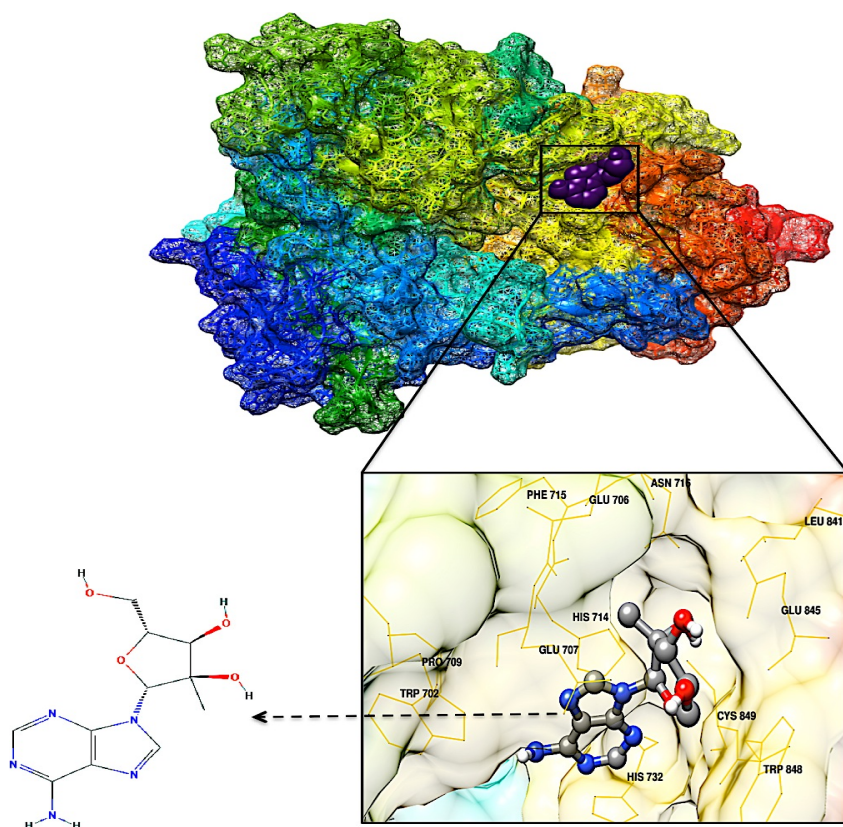


Figure 4.6: Docked Conformation of 2'-C-Methyladenosine with ZIKV NS5 RdRp (Binding affinity: -6.3kcal/mol)

Further information, including extensive procedures, can be found in our previous publications on structure-based enzymatic drug design (58,75). We believe that the robust computational tools implemented in the route map will provide a fundamental platform in the development of inhibitors against multiple ZIKV target molecules.

6. Proposed computational software that can be used in ZIKV drug design and discovery

The software available for techniques used in drug design have simplified the development of inhibitors allowing for specific binding to a target molecule, thus, decreasing its biological adverse effects (76). There are various types of software available in structure-based drug design, allowing for faster and more comprehensive research into ZIKV inhibitors (Table 4).

Table 4.4: Proposed computational software used in ZIKV drug design

COMPUTATIONAL METHOD	SOFTWARE AVAILABLE	SOFTWARE TO BE UTILIZED IN STUDY
Homology Modeling	<u>Sequence alignment-</u> Insight, Prime, Profit, LOOK, ICM, Sybyl, CLUSTALW <u>Model construction-</u> DS Modeller, Prime, LOOK, ICM, Sybyl, MODELLER, MOE, SWISS-MODEL, RaptorX, LOMETS, Phyre, I-Tasser	<u>Sequence alignment-</u> CLUSTALW (77) <u>Model construction-</u> MODELLER (78)
Active Binding Site Determination	CASTp, POOL, PASS, Pocket-Finder, 3DLigandSite, LIGSITE, metaPocket, FINDSITE, Site-Hound	metaPocket (79) POOL (80) Site-Hound(71)
Molecular Graphic systems	Avogadro, Chemlab, Athena, Maestro, Jmol, PyMOL, UCSF chimera, VMD, Vimol, Webmol, Zeus	UCSF chimera (12)
Virtual Screening databases	PubChem, MMsINC, ZINC, ZincPharmer, 4SC discovery, Therapeutic target database, Drug Bank, ChemSpider, ChEMBL	ZINC (81) ZincPharmer (82)
Docking Software	PyRx, Autodock Vina, Dock Blaster, Vis3d, Schrodinger, GOLD, Libdock, FlexX, Glide, Fred, ICM	Autodock Vina (83)

Molecular dynamic simulation software	Gromacs, Amber, CHARM, Gromos, ADF, Desmond, NWChem	Amber (84)
---------------------------------------	---	------------

7. Conclusion

The Future of the ZIKV pandemic is uncertain and thus new, accelerated techniques are necessary to assist the medical and scientific community in the identification and validation of inhibitors to this global threat. The chemoinformatics discussed in this paper will not only in the identification and design of potential ZIKV inhibitors but also in parallel, but may assist in the early analysis of potential biological mutations that may occur due to the rapid international transmission of this flavivirus.

Acknowledgements

The authors acknowledge the College of Health Sciences, UKZN, and the National Research Foundation for their financial support.

Conflicts of interest

Authors declare no potential conflicts of interest.

References

1. Faye O, Freire CCM, Iamarino A, Faye O, de Oliveira JVC, Diallo M, *et al.* Molecular Evolution of Zika Virus during Its Emergence in the 20th Century. *PLoS Negl Trop Dis.* 2014;8(1):1–10.
2. Bhakat S, Karubiu W, Jayaprakash V, Soliman MES. A perspective on targeting non-structural proteins to combat neglected tropical diseases: Dengue, West Nile and Chikungunya viruses. *Eur J Med Chem* [Internet]. Elsevier Masson SAS; 2014;87(2014):677–702.
3. Chen LH, Hamer DH. Zika Virus: Rapid Spread in the Western Hemisphere. *Ann Intern Med.* 2016:1–3.
4. Campos GS, Bandeira AC, Sardi SI. Zika virus outbreak, Bahia, Brazil. *Emerg Infect Dis.* 2015;21(10):1885–6.
5. Mahfuz M, Khan A, Mahmud H Al, Hasan M, Parvin A, Rahman N, *et al.* Indian Journal of Pharmaceutical and Biological Research (IJPBR) In Silico Modeling and Immunoinformatics Probing Disclose the Epitope Based Peptide Vaccine Against Zika Virus Envelope Glycoprotein. *Indian J Pharm Biol Res.* 2014;2(4):44–57.
6. Roa M. Zika virus outbreak: Reproductive health and rights in Latin America. *Lancet.* 2016;387(10021):843.
7. WHO. Zika virus, Microcephaly and Guillain-Barré syndrome. 2016.
8. Musso D, Roche C, Robin E, Nhan T, Teissier A, Cao-Lormeau VM. Potential sexual transmission of zika virus. *Emerg Infect Dis.* 2015;21(2):359–61.
9. Ekins S, Mietchen D, Coffee M, Stratton T, Freundlich J, Freitas-Junior L, *et al.* Open drug discovery for the Zika virus. *F1000Research* [Internet]. 2016;5(0):150.
10. Singh RK, Dhama K, Malik YS, Ramakrishnan MA, Karthik K, Tiwari R, *et al.* Zika Virus - Emergence, evolution, pathology, diagnosis and control: current global scenario and future perspectives - A comprehensive review. *Vet Q* [Internet]. 2016;2176(July):1–43.
11. Lissauer D, Smit E, Kilby MD. Zika virus and pregnancy. *An Int J Obstet Gynaecol* [Internet]. 2016;1258–63.
12. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* [Internet]. 2004;25(13):1605–12.
13. Panchaud A, Stojanov M, Ammerdorffer A, Vouga M. Emerging Role of Zika Virus in Adverse Fetal and Neonatal Outcomes. *Clin Microbiol Rev.* 2016;29(3):659–94.
14. Chibueze EC, Tirado V, Olukunmi O. Zika virus infection in pregnancy : a systematic

review of disease course and complications. Bull World Health Organ. 2016;(June).

15. Mlakar J, Korva M, Tul N, Popović M, Poljšak-Prijatelj M, Mraz J, *et al.* Zika Virus Associated with Microcephaly. N Engl J Med [Internet]. 2016;374(10):951–8.
16. Jean Marie Turmel PA, Bruno Hubert YMV, Maquart M, Guillou-Guillemette H Le, Art IL-G. Late sexual transmission of Zika virus related to. Lancet. 2016;6736(16):30775.
17. Oliveira Melo AS, Malinger G, Ximenes R, Szejnfeld PO, Alves Sampaio S, Bispo De Filippis AM. Zika virus intrauterine infection causes fetal brain abnormality and microcephaly: Tip of the iceberg? Ultrasound Obstet Gynecol. 2016;47(1):6–7.
18. Plourde AR, Bloch EM. A Literature Review of Zika Virus. Emerg Infect Dis [Internet]. 2016;22(7):1–15.
19. Calvet G, Aguiarv RS, Melo a S, Sampaio S a, de Filippis I, Fabri A, *et al.* Case Report of detection of Zika virus genome in amniotic fluid of affected fetuses: association with microcephaly outbreak in Brazil. Lancet Infect Dis. 2016;3099(16):1–8.
20. Petersen LR, Jamieson DJ, Powers AM, Honein MA. Zika Virus. N Engl J Med [Internet]. 2016;374(16):1552–63.
21. Waggoner JJ, Pinsky BA. Zika Virus: Diagnostics for an Emerging Pandemic Threat. J Clin Microbiol [Internet]. 2016;54(4):860–7.
22. Gourinat AC, O, Connor O, Calvez E, Goarant C, Dupont-Rouzeyrol M. Detection of zika virus in urine. Emerg Infect Dis. 2015;21(1):84–6.
23. Charrel RN, Leparac-Goffart I, Pas S, de Lamballerie X KM& RC. State of knowledge on Zika virus for an adequate laboratory response. Bull World Health Organ [Internet]. 2016;(February):1–29.
24. Haddow AD, Schuh AJ, Yasuda CY, Kasper MR, Heang V, Huy R, *et al.* Genetic characterization of zika virus strains: Geographic expansion of the asian lineage. PLoS Negl Trop Dis. 2012;6(2):1–7.
25. Hayes EB. Zika virus outside Africa. Emerg Infect Dis. 2009;15(9):1347–50.
26. DICK GW a, KITCHEN SF, HADDOW AJ. Zika virus. I. Isolations and serological specificity. Trans R Soc Trop Med Hyg [Internet]. 1952;46(5):509–20. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12995440>
27. Duffy MR, Chen T-H, Hancock WT, Powers AM, Kool JL, Lanciotti RS, *et al.* Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med. 2009;360(24):2536–43.
28. Malone RW, Homan J, Callahan M V., Glasspool-Malone J, Damodaran L, Schneider ADB, *et al.* Zika Virus: Medical Countermeasure Development Challenges. PLoS Negl Trop Dis. 2016;10(3):1–26.

29. Basarab M, Bowman C, Aarons EJ, Cropley I. **Zika virus**. Bmj [Internet]. 2016;1049(February):i1049.
30. Saiz J-C, Vázquez-Calvo Á, Blázquez AB, Merino-Ramos T, Escribano-Romero E, Martín-Acebes MA. Zika Virus: the Latest Newcomer. Front Microbiol [Internet]. 2016;7(April):496.
31. Ekins S, Liebler J, Neves BJ, Lewis WG, Coffee M, Bienstock R, *et al.* Illustrating and homology modeling the proteins of the Zika virus. F1000Research [Internet]. 2016;5(0):275.
32. Hughes JP, Rees SS, Kalindjian SB, Philpott KL. Principles of early drug discovery. Br J Pharmacol. 2011;162(6):1239–49.
33. Noble CG, Chen YL, Dong H, Gu F, Lim SP, Schul W, *et al.* Strategies for development of dengue virus inhibitors. Antiviral Res [Internet]. Elsevier B.V.; 2010;85(3):450–62.
34. Perera R, Khaliq M, Kuhn RJ. Closing the door on flaviviruses: Entry as a target for antiviral drug design. Antiviral Res. 2008;80(1):11–22.
35. Chappell KJ, Stoermer MJ, Fairlie DP, Young PR. Insights to substrate binding and processing by West Nile Virus NS3 protease through combined modeling, protease mutagenesis, and kinetic studies. J Biol Chem. 2006;281(50):38448–58.
36. Agnihotri S, Narula R, Joshi K, Rana S, Singh M. In silico modeling of ligand molecule for non structural 3 (NS3) protein target of flaviviruses. Bioinformation [Internet]. 2012;8(3):123–7.
37. Tian H, Ji X, Yang X, Xie W, Yang K, Chen C, *et al.* The crystal structure of Zika virus helicase: basis for antiviral drug design. Protein Cell [Internet]. 2016;7(6):450–4.
38. Song H, Qi J, Haywood J, Shi Y, Gao GF. Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses. Nat Struct Mol Biol [Internet]. 2016;(May):1–4.
39. Youn S, Ambrose RL, Mackenzie JM, Diamond MS. Non-structural protein-1 is required for West Nile virus replication complex formation and viral RNA synthesis. Virol J [Internet]. 2013;10:339.
40. Beatty PR, Puerta-Guardo H, Killingbeck SS, Glasner DR, Hopkins K, Harris E. Dengue virus NS1 triggers endothelial permeability and vascular leak that is prevented by NS1 vaccination. Sci Transl Med [Internet]. 2015;7(304):304ra141.
41. Stahla-Beek HJ, April DG, Saeedi BJ, Hannah a. M, Keenan SM, Geiss BJ. Identification of a Novel Antiviral Inhibitor of the Flavivirus Guanylyltransferase Enzyme. J Virol. 2012;86(16):8730–9.
42. Sirohi D, Chen Z, Sun L, Klose T, Pierson TC, Rossmann MG, *et al.* The 3.8 Å resolution cryo-EM structure of Zika virus. Science [Internet]. 2016;352(6284):467–70.
43. Courageot M, Frenkiel M, Santos D Dos, Deubel V, Desprès P, Duarte C, *et al.* α -

Glucosidase Inhibitors Reduce Dengue Virus Production by Affecting the Initial Steps of Virion Morphogenesis in the Endoplasmic Reticulum α -Glucosidase Inhibitors Reduce Dengue Virus Production by Affecting the Initial Steps of Virion Morphogenesis in. J Virol. 2000;74(1):564–72.

44. Hamel R, Dejarnac O, Wichit S, Ekchariyawat P, Neyret A, Luplertlop N, *et al.* Biology of Zika Virus Infection in Human Skin Cells. J Virol [Internet]. 2015;89(17):8880–96.
45. Noppakunmongkolchai W, Poyomtip T, Jittawuttipoka T, Luplertlop N, Sakuntabhai A, Chimnaronk S, *et al.* Inhibition of protein kinase C promotes dengue virus replication. Virol J [Internet]. Virology Journal; 2016;13(1):35.
46. Hirsch AJ. in Viral Replication. Future Med Chem. 2010;5(2):303–11.
47. Zhang R, Miner JJ, Gorman MJ, Rausch K, Ramage H, White JP, *et al.* A CRISPR screen defines a signal peptide processing pathway required by flaviviruses [Internet]. Nature. 2016.
48. Bash E. Design and Prediction of Potential RNAi (siRNA) Molecules for 3' UTR PTGS of different strains of Zika Virus: A Computational Approach. Nat Sci. 2015;1(2):37–50.
49. Kostyuchenko VA, Lim EXY, Zhang S, Fibriansah G, Ng T-S, Ooi JSG, *et al.* Structure of the thermally stable Zika virus. Nature [Internet]. Nature Publishing Group; 2016;533(7603):425–8.
50. Dai L, Song J, Lu X, Deng Y-Q, Musyoki AM, Cheng H, *et al.* Structures of the Zika Virus Envelope Protein and Its Complex with a Flavivirus Broadly Protective Antibody. Cell Host Microbe [Internet]. Elsevier Inc.; 2016;19(5):696–704.
51. Barba-Spaeth G, Dejnirattisai W, Rouvinski A, Vaney M-C, Medits I, Sharma A, *et al.* Structural basis of potent Zika–dengue virus antibody cross-neutralization. Nature [Internet]. Nature Publishing Group; 2016;1–23.
52. Weltman JK. Medical Microbiology & Diagnosis An Immuno-Bioinformatic Analysis of Zika virus (ZIKV) Envelope E Protein. J Med Microbiology Diagnosis. 2016;5(2).
53. Mandal S, Moudgil M, Mandal SK. Rational drug design. Eur J Pharmacol [Internet]. Elsevier B.V.; 2009;625(1-3):90–100.
54. Song CM, Lim SJ, Tong JC. Recent advances in computer-aided drug design. Brief Bioinform. 2009;10(5):579–91.
55. Bamborough P, Cohen FE. Modeling protein-ligand complexes. Curr Opin Struct Biol [Internet]. 1996;6(2):236–41.
56. Anderson AC. The Process of Structure-Based Drug Design. Chem Biol. 2003;10(9):787–97.
57. Huang HJ, Yu HW, Chen CY, Hsu CH, Chen HY, Lee KJ, *et al.* Current developments of computer-aided drug design. J Taiwan Inst Chem Eng [Internet]. Taiwan Institute of

Chemical Engineers; 2010;41(6):623–35.

58. Honarparvar B, Govender T, Maguire GEM, Soliman MES, Kruger HG. Integrated Approach to Structure-Based Enzymatic Drug Design : Molecular Modeling , Spectroscopy , and Experimental Bioactivity. *Chem Rev*. 2014;114:493–537.
59. Hooft RW, Sander C, Vriend G. Objectively judging the quality of a protein structure from a Ramachandran plot. *Comput Appl Biosci*. 1997;13(4):425–30.
60. Chetty S, Soliman MES. Possible allosteric binding site on Gyrase B , a key target for novel anti-TB drugs : homology modelling and binding site identification using molecular dynamics simulation and binding free energy calculations. *Med Chem Res [Internet]*. Springer US; 2015;24(2015):2055–74.
61. Saberi Fathi SM, Tuszynski J a. A simple method for finding a protein's ligand-binding pockets. *BMC Struct Biol [Internet]*. 2014;14(18):1–9.
62. Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. *Curr Top Med Chem [Internet]*. 2014;14(16):1923–38.
63. Ramesh M, Vepuri SB, Oosthuizen F, Soliman ME. Adenosine Monophosphate-Activated Protein Kinase (AMPK) as a Diverse Therapeutic Target: A Computational Perspective. *Appl Biochem Biotechnol [Internet]*. 2015;1–21.
64. Reddy SS, Pati PP, Kumar PP, Pradeep H, Sastry NN. Virtual screening in drug discovery -- a computational perspective. *Curr Protein Pept Sci*. 2007;8(4):329–51.
65. Sliwoski G, Kothiwale S, Meiler J, Lowe EW. Computational methods in drug discovery. *Pharmacol Rev [Internet]*. 2014;66(1):334–95.
66. Kumalo HM, Soliman ME. Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery: A Case Study on the Identification of Novel β -Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents. *Cell Mol Bioeng*. 2015;9(1):175–89.
67. Delvecchio R, Higa LM, Pezzuto P, Valadao AL, Garcez PP, Monteiro FL, *et al*. Chloroquine inhibits Zika Virus infection in different cellular models. *bioRxiv [Internet]*. 2016;051268. Available from: <http://biorxiv.org/content/early/2016/05/02/051268.abstract>
68. Dowall SD, Graham VA, Rayner E, Atkinson B, Hall G, Watson RJ, *et al*. A Susceptible Mouse Model for Zika Virus Infection. *PLoS Negl Trop Dis [Internet]*. 2016;10(5):e0004658.
69. Larocca RA, Abbink P, Peron JPS, Zanotto PM de A, Iampietro MJ, Badamchi-Zadeh A, *et al*. Vaccine protection against Zika virus from Brazil. *Nature [Internet]*.
70. Maharaj Y, Soliman MES. Identification of novel gyrase b inhibitors as potential Anti-TB drugs: Homology modelling, hybrid virtual screening and molecular dynamics

simulations. *Chem Biol Drug Des.* 2013;82(2):205–15.

71. Hernandez M, Ghersi D, Sanchez R. SITEHOUND-web: A server for ligand binding site identification in protein structures. *Nucleic Acids Res.* 2009;37(SUPPL. 2):413–6.
72. Tambunan USF, Zahroh H, Utomo BB, Parikesit AA. Screening of commercial cyclic peptide as inhibitor NS5 methyltransferase of Dengue virus through Molecular Docking and Molecular Dynamics Simulation. *Bioinformation [Internet].* 2014;10(1):23–7.
73. Niyomrattanakit P, Chen YL, Dong H, Yin Z, Qing M, Glickman JF, *et al.* Inhibition of dengue virus polymerase by blocking of the RNA tunnel. *J Virol [Internet].* 2010;84(11):5678–86.
74. Eyer L, Nencka R, Huvarova I, Palus M, Alves M, Gould EA, *et al.* Nucleoside inhibitors of Zika virus. *J Infect Dis.* 2016;27(May):1–13.
75. Moonsamy S, Bhakat S, Soliman MES. Dynamic features of apo and bound HIV-Nef protein reveal the anti-HIV dimerization inhibition mechanism. *J Recept Signal Transduct [Internet].* 2014;9893:1–11.
76. Liao C, Sitzmann M, Pugliese A, Nicklaus MC. Software and resources for computational medicinal chemistry. *Future Med Chem.* 2011;3(8):1057–85.
77. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics [Internet].* 2002;Chapter 2(2.3):1–22.
78. Eswar N, Webb B, Marti-Renom M a, Madhusudhan MS, Eramian D, Shen MY, *et al.* Comparative protein structure modeling using Modeller [Internet]. *Curr Protoc Bioinformatics.* 2006. p. 1–47.
79. Somarowthu S, Yang H, Hildebrand DGC, Ondrechen MJ. High-performance prediction of functional residues in proteins with machine learning and computed input features. *Biopolymers.* 2011;95(6):390–400.
80. Huang B. MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS.* 2009;13(4):325–30.
81. Irwin JJ, Shoichet BK. ZINC - A free database of commercially available compounds for virtual screening. *J Chem Inf Model.* 2005;45(1):177–82.
82. Koes DR, Camacho CJ. ZINCPharmer: Pharmacophore search of the ZINC database. *Nucleic Acids Res.* 2012;40(W1):409–14.
83. Trott O, Olson AJ. AutoDock Vina. *J Comput Chem.* 2010;31:445–61.
84. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM, *et al.* The Amber biomolecular simulation programs. *Journal of Computational Chemistry.* 2005. p. 1668–88.

CHAPTER 5

Zika Virus NS5 Protein Potential Inhibitors: An Enhanced *In silico* Approach in Drug Discovery

Pritika Ramharack^A and Mahmoud E. S. Soliman^{A*}

^AMolecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

Corresponding Author: Mahmoud E.S. Soliman

1. Dean and Head of School of Health Sciences, Full Professor: Pharmaceutical Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.
2. Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt.
3. College of Pharmacy and Pharmaceutical Sciences, Florida Agricultural and Mechanical University, FAMU, Tallahassee, Florida 32307, USA.

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Acknowledgements

The authors acknowledge the College of Health Sciences, UKZN, and the National Research Foundation for their financial support and the Center for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

Abstract

The re-emerging Zika virus is an arthropod-borne virus that has been described to have explosive potential as a worldwide pandemic. The initial transmission of the virus was through a mosquito vector, however, evolving modes of transmission has allowed the spread of the disease over continents. The virus has already been linked to irreversible chronic central nervous system conditions. The concerns of the scientific and clinical community are the consequences of Zika viral mutations, thus suggesting the urgent need for viral inhibitors. There have been large strides in vaccine development against the virus but there are still no FDA approved drugs available. Rapid rational drug design and discovery research is fundamental in the production of potent inhibitors against the virus that will not just mask the virus, but destroy it completely. *In silico* drug design allows for this prompt screening of potential leads, thus decreasing the consumption of precious time and resources. This study demonstrates an optimized and proven screening technique in the discovery of two potential small molecule inhibitors of Zika virus Methyltransferase and RNA dependent RNA polymerase. This *in silico* “per-residue energy decomposition pharmacophore” virtual screening approach will be critical in aiding scientists in the discovery of not only effective inhibitors of Zika viral targets, but also a wide range of anti-viral agents.

Keywords:

Zika virus per-residue decomposition based pharmacophore, virtual screening, NS5 protein potential inhibitors, binding free energy, molecular dynamic simulations.

Introduction:

Zika virus (ZIKV) is an arthropod-borne virus that has been described to have potential as a worldwide pandemic (Troncoso, 2016). The virus is a member of the *spodweni* serocomplex of the *flavivirus* genus and was first discovered in 1947 by its isolation from the Rhesus 766 monkey in Uganda (Faye *et al.*, 2014; Haddow *et al.*, 2012). Sporadic cases of the virus have been reported in countries such as Uganda, Tanzania, Egypt, Gabon, and in parts of Asia including India and Indonesia, with the most devastating epidemic occurring in Brazil in 2015 (Campos, Bandeira, & Sardi, 2015; Mahfuz *et al.*, 2015). As of June 2016, eleven countries had reported Central Nervous System (CNS) malformations potentially linked to ZIKV. During 2015 and early 2016, eight countries had reported cases of Guillain-Barré syndrome (GBS), where laboratory testing confirmed ZIKV infection was found in a number of GBS cases (WHO, 2016).

Transmission of the virus was thought to be only via the *Aedes* mosquito vector but studies during 2016, have evidenced congenital, perinatal and sexual transmission (Singh *et al.*, 2016; Turmel, Hubert, Maquart, Guillou-Guillemette, & Leparç-Goff, 2016). The virus triggers febrile like influenza-conditions in the host, including swollen lymph nodes, skin rashes and joint pains (Brito, 2016; Ekins *et al.*, 2016; Shapshak, Sinnott, Somboonwit, & Kuhn, 2015). The concerns of the scientific community involve the dramatic increase in ZIKV-related central nervous system (CNS) disorders including neonatal-microcephaly and Guillain-Barré Syndrome (Lissauer, Smit, & Kilby, 2016; Panchaud, Stojanov, Ammerdorffer, & Vouga, 2016; Roa, 2016). Complications associated with prenatal infection encompass fetal growth restriction, neurological and ocular abnormalities, intracranial calcification and in some cases perinatal death or stillbirth (Chibueze, Tirado, & Olukunmi, 2016; Singh *et al.*, 2016).

The virus is able to enter a host via receptor-mediated endocytosis, followed by fusion from within the endosomal cell compartment (Mahfuz *et al.*, 2014). The enveloped virus comprises of an 11 kilo base, single-stranded positive sense RNA genome which consists of 10,794 nucleotides encoding 3,419 amino acids (Hayes, 2009). The open reading frame (ORF) of the 5' and 3' untranslated region encodes a polyprotein cleaved into three structural proteins being the capsid, premembrane/membrane, and envelope. Seven non-structural proteins may also be found in this assembly, namely, NS1, NS2A, NS2B, NS3, NS4A, 2K, NS4B, and NS5 (largest viral protein) (Haddow *et al.*, 2012).

Being the largest and most imperative protein in the genome replication and RNA capping of ZIKV, NS5 presents as a novel antiviral target (Tambunan, Zahroh, Utomo, & Parikesit, 2014). The protein consists of three domains: a Methyltransferase (MTase) domain at residues 1-262 of its N-terminal, an RNA dependent RNA polymerase (RdRp) at residues 273-903 of its C-terminal and an inter-domain region at residues 263-272 (Zou *et al.*, 2014).

The MTase domain belongs to the family of S-Adenosyl Methionine (SAM)- dependent enzymes, containing a SAM-dependent MTase fold comprising of an $\alpha/\beta/\alpha$ structure (Zou *et al.*, 2014). The MTase domain is one of the key targets in drug design as the enzyme performs nucleoside-2'O and N-7 methylation of the viral RNA cap which is essential in the replication of the virus (Egloff, Benarroch, Selisko, Romette, & Canard, 2002). Upon the completion of methylation, SAM is converted to S-Adenosyl Homocysteine (SAH) and gets released from the MTase domain (Brecher *et al.*, 2015). Inhibition of MTase will be detrimental to the progression of ZIKV.

The conserved RdRp domain allows for the initiation of RNA synthesis, generating both plus and

minus strand RNAs. As with most polymerases, the structure of the enzyme resembles a shape analogous to a right hand with a finger, thumb and palm region (Papageorgiou *et al.*, 2014; Shanmugam, Velmurugan, & Gromiha, 2015). The human body does not contain an RdRp enzyme or analogues of it, thus inhibitors may not cause severe toxic effects, making it an optimal target in drug design (Shanmugam *et al.*, 2015).

To date, no anti-ZIKV drugs are clinically available, thus, new research methods are being developed with the purpose of identifying target molecules. Recent research has found that ZIKV targets neuronal cells (Millichap, 2016; Miner & Diamond, 2016; Mlakar *et al.*, 2016; Tang *et al.*, 2016). Consequently, any new drugs that may be discovered will have to pass through the blood-brain-barrier. Molecular modeling and computational methods are important tools in the development of novel inhibitors of ZIKV (Ekins *et al.*, 2016; Ramharack & Soliman, 2016). A number of inhibitors of the *flavivirus* NS5 protein have been discovered via virtual screening and computational analysis (Brecher *et al.*, 2015; Idrus, Tambunan, & Zubaidi, 2012; Lim & Shi, 2013). Structure-based virtual screening (SBVS) identifies energetically advantageous binding affinities of ligands into a target's active binding site. This allows for new insights on the nature of the active site and the protein-ligand interactions (Kumalo & Soliman, 2015). The method identifies selective molecules from an extensive library of compounds to dock within a target's active site. Although scoring techniques are used when molecules are docked to the target, literature shows that a large number of final hits are generated, as the compounds docked may be in various geometric poses (Kroemer, 2007). Ligand-based virtual screening generates libraries of compounds based on a known compound or compounds and its illustrative interactions with a particular target (Cele, Muthusamy, & Soliman, 2016).

In an attempt to develop pharmacophore based modeling, we previously presented a Per Residue Energy Decomposition (PRED) protocol where candidates for SBVS were chosen on the position

of 3D moieties with an experimentally known compound, thus creating a pharmacophore model based on highly contributing amino acid residues to the bound inhibitor. This approach is based on interactions that occur at a molecular level, including charge, hydrophobic interactions and hydrogen bonding (Cele *et al.*, 2016). The highly contributing residues are identified based on free energy footprints from molecular dynamic and thermodynamic calculations (Cele *et al.*, 2016; Kumalo & Soliman, 2015; Soliman, 2013). This proves to be an incredibly concise method, rather than “shooting in the dark” with millions of available small molecules.

In our previous work, we created a possible homology model of the NS5 protein containing both MTase and RdRp domains (Ramharack & Soliman, 2016). Due to the indeterminateness surrounding the ZIKV NS5 protein and potential inhibitors, we will compare our top hits against known inhibitors of the *flavivirus* NS5 protein.

This study will implement the above-mentioned PRED pharmacophore technique in the discovery of potential ZIKV NS5 protein inhibitors, thus aiding medicinal chemists in the synthesis of possible drug candidates.

2. Computational Methods

A route map to PRED-based pharmacophore virtual screening approach is depicted in Figure 5.1.

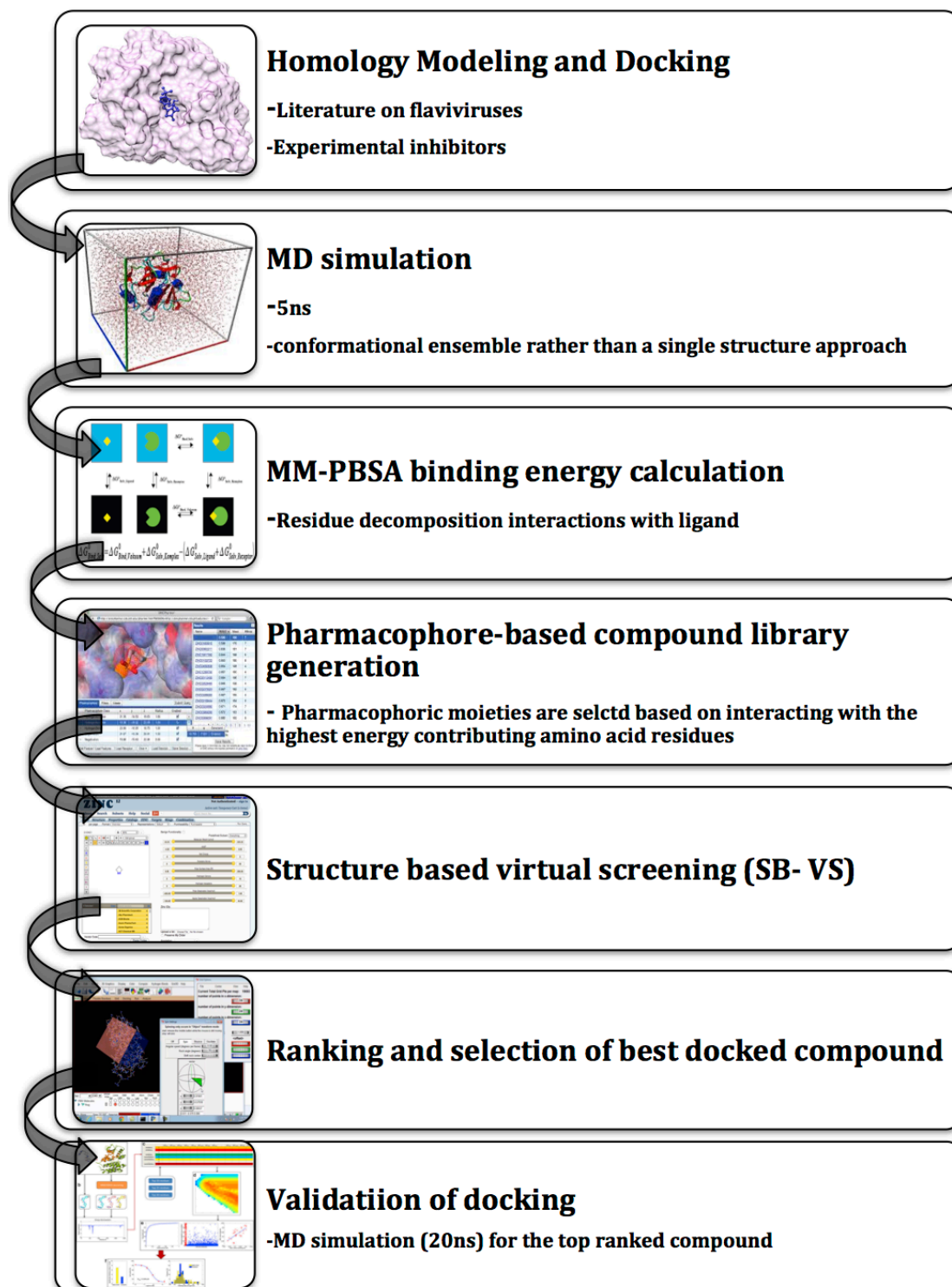


Figure 5.1: A per residue energy decomposition-based approach outline applied in the study.

2.1 Homology Modeling and Identification of active binding sites of NS5

In our previous article and during the current study, due to the absence of a crystal structure of ZIKV NS5 protein, a homology model was created using the protein sequence obtained from NCBI (Accession number: YP_009227205) (Ramharack & Soliman, 2016). The templates for sequence alignment were identified from NCBI using BLASTp (accessed on 05 March 2016) (Madden, Tatusov, & Zhang, 1996) to find suitable templates, from RCSB protein databank (Berman *et al.*, 2002), for homology modeling. Based on the criteria of identity score, e-value and query cover accuracy (Table 1), the NS5 protein was modeled by using four crystal structures of *flavivirus* enzymes as templates: Chain A of full-length Japanese Encephalitis Virus NS5 (PDB Code: 4K6M_A); Chain A of full-length NS5 from Dengue virus Type 3 (PDB Code: 5CCV_A); Chain A of RNA Dependent RNA polymerase domain from Nile West Virus (PDB Code: 2HFZ_A); and Chain A of Dengue Serotype 3 RNA-dependent RNA polymerase bound to Nitd-107 (PDB Code: 3VWS_A).

Table 5.1: Criteria summary of chosen templates used in Building the ZIKV NS5 homology model.

Template PDB Code:	Query Cover (%)	Structural Identity (%)	E-Value
4K6M_A	97	69	0
5CCV_A	98	67	0
2HFZ_A	67	72	0
3VWS_A	68	69	0

Homology modeling was performed using the Modeller Software version 9.1 (Eswar *et al.*, 2006) add-on in chimera (Yang *et al.*, 2012), in which all three templates were selected to build the model. Multiple sequence alignment was performed using the CLUSTALW server (Sievers *et al.*, 2011), where Chain A of the full-length Japanese Encephalitis Virus NS5 was evidenced to have the best template with the highest identity score (Figure S1). The sequence of the target protein was uploaded to PSIPRED V3.3 (Buchan, Minneci, Nugent, Bryson, & Jones, 2013) in order to obtain a predicted 2D secondary structure of the enzyme. Comparing the homolog to the predicted 2D structure and assessment of the bond angles and torsional strain validated the homology model. A Ramachandran plot for the analyses of bond angles and torsional strain was generated using Maestro (Schrodinger). MolProbity (Chen *et al.*, 2010) results showed 97.2% of all residues were in the favored regions and 99.2% of all residues were in the allowed regions, which left a list of 7 outliers. The active-site residues were determined using Chimera Multi-align Viewer and validated using the SiteHound-web program (Hernandez, Ghersi, & Sanchez, 2009). The list shows that none of the active-site residues are part of these outliers. All results can be found in our previous article and supplementary material (Figure S5) (Ramharack & Soliman, 2016). After completion of the study, the crystal structure of the ZIKV NS5 protein was released. To validate the homology model of the NS5, it was superimposed with the newly released crystal structure (PDB code: 5TFR), showing their structural similarity and validating the model's use for subsequent analysis (Figure 5.2).

2.2 System Preparation

The NS5 modeled structure was separated into two domains, being the Methyltransferase of the C-terminal and the RNA-Dependent RNA polymerase of the N-terminal. Experimental drug inhibitors of *flaviviruses* were chosen to dock within each domain's active site.

2.3 Molecular Docking of Experimental Flaviviruses

Docking of the compounds were conducted using the AutoDock Vina (Trott & Olson, 2010) software. The procedure was run using the software default settings. The grid box used to define the screening site was elucidated using the AutoDock Vina functionality built into Chimera (Eric F. Pettersen *et al.*, 2004). The gridbox size and center parameters for the MTase were x(54,-63.23), y(80,56.72) and z(54,10.22), respectively and the RdRp gridbox dimensions were x(40, -9.69), y(38, 20.41), z(40,16.50). AutoDock Vina generated results in the pdbqt format and the optimal geometric conformation having the best binding energy was selected from the ViewDock feature and saved in complex with the reference enzyme. The enzyme and ligand for each system was prepared using Chimera (Yang *et al.*, 2012) and MMV molecular modeling suites (Kusumaningrum *et al.*, 2014) and subsequently subjected to molecular dynamic simulations.

2.4 Molecular Dynamic (MD) Simulations

The MD simulation was performed using the GPU version of the PMEMD engine provided with the Amber 14 package. The FF14SB force field of the Amber package (Nair & Miners, 2014) was used to describe the complex.

ANTECHAMBER (Wang, Wang, Kollman, & Case, 2006) was used to generate atomic partial charges for the ligands by utilizing the Restrained Electrostatic Potential (RESP) and the General Amber Force Field (GAFF) procedures. The Leap module of Amber 14 allowed for addition of hydrogen atoms to the systems as well as Na⁺ and Cl⁻ counter ions for neutralization.

The system was suspended implicitly within an orthorhombic box of TIP3P water molecules such

that all atoms were within 8 Å of any box edge.

An initial minimization of 2000 steps was carried out with an applied restraint potential of 500 kcal/mol Å² for both complexes. An additional full minimization of 1000 steps was further carried out by conjugate gradients algorithm without restrain.

A gradual heating MD simulation from 0 K to 300 K was executed for 50 ps, such that the system maintained a fixed number of atoms and fixed volume, i.e., a canonical ensemble (NVT). The solutes within the system are imposed with a potential harmonic restraint of 10 kcal/mol Å² and collision frequency of 1.0 ps⁻¹. Following heating, an equilibration estimating 500 ps of the each system was conducted, the operating temperature was kept constant at 300 K. Additional features such as a number of atoms and pressure were also kept constant mimicking an isobaric-isothermal ensemble (NPT). The systems pressure was maintained at 1 bar using the Berendsen barostat.

The total time for the MD simulation conducted was 5 ns. In each simulation the SHAKE algorithm was employed to constrict the bonds of hydrogen atoms. The time step of each simulation was 2 fs and an SPFP precision model was used. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of 1 bar maintained by the Berendsen barostat, a pressure-coupling constant of 2 ps, a temperature of 300 K and Langevin thermostat with collision frequency of 1.0 ps⁻¹.

Coordinates were saved every 1 ps and the trajectories were analyzed every 1 ps using the PTRAJ module employed in Amber14.

2.5 Binding Free energy Calculations

To estimate the binding affinities of each system, the binding free energies were calculated using

the Molecular Mechanics/GB Surface Area method (MM/GBSA) (Genheden & Ryde, 2015). Binding free energies were averaged over 5000 snapshots extracted from the 5 ns trajectory. The free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

$$(1) \Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}}$$

$$(2) \Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS$$

$$(3) E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}}$$

$$(4) G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}}$$

$$(5) G_{\text{SA}} = \gamma \text{SASA}$$

The term E_{gas} denotes the gas-phase energy, which consist of the internal energy E_{int} ; Coulomb energy E_{ele} , and the van der Waals energies E_{vdw} . The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol} , was estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G . The non-polar solvation energy, G_{SA} .

G_{SA} , was determined from the solvent accessible surface area (SASA), using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB} , contribution was estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature respectively.

To obtain the contribution of each residue to the total binding free energy profile between the inhibitors Ribavirin and BG323 with RdRp and MTase respectively, per-residue free energy

decomposition was carried out at the atomic level for imperative residues using the MM/GBSA method in Amber 14.

2.6 Pharmacophore Model Creation and Library Generation

The inhibitors Ribavirin and BG323 were first simulated at the active site of RdRp and MTase respectively, for 5 ns, to create the bound conformation of the ligands. Both these compounds have experimentally exhibited ZIKV inhibition in *in vitro* and *in vivo* models (Mumtaz, van Kampen, Reusken, Boucher, & Koopmans, 2016; Sweeney *et al.*, 2015; Zmurko *et al.*, 2016). Per-residue energy decomposition analysis was used to determine the amino acids that contribute the most towards ligand binding. The pharmacophoric moieties that interacted with the highly contributing residues were then chosen to construct our model. The model was then added to ZincPharmer (Koes & Carlos, 2012), with specific selection criteria (molecular weight of <500 Da, rotatable bonds <6, hydrogen bond donors<5 and hydrogen bond acceptors<10), to screen the ZINC database (Irwin & Shoichet, 2005). Lipinski's rule of five and toxicity (ADMET) properties were used as filters to remove nondrug-like hits (Lipinski, Lombardo, Dominy, & Feeney, 2012).

2.7 Structure-based Virtual Screening

The drug-like hits identified using our protocol were subjected to structure-based virtual screening. Docking was carried out to differentiate between ligands based on the molecules' geometric characteristics that allow it to bind to the enzyme's active site (Forli *et al.*, 2016). The Docking calculations were performed using Autodock Vina (Trott & Olson, 2010). During docking, Gasteiger partial charges were assigned and the Autodock atom types were defined using the Autodock Graphical user interface supplied by MGL tools (Sanner, Olson, & Spehner,

1996). The docked conformations were generated using the Lamarckian Genetic Algorithm (Morris & Huey, 2009). The Raccoon software was used to convert the files into a compatible pdbqt format required for docking. The gridbox was defined using Autodock Vina. The calculation reports for each ligand conformation in its respective complex were analyzed to obtain affinity energy (kcal/mol). During the docking process, a maximum of 50 conformers was considered for each compound. After screening, molecular docking and filtering, the ligand with the highest affinity towards the agonist was selected from the library.

2.8 Validation of Docking Approach

Previous experiences have verified that docking may result in the best geometric conformation of the docked complex, however, short molecular dynamic simulations may not be able to maintain the stability of the complex and thus lead to the molecules being disorientated. Thus, to validate the approach applied in this study, the most favorable Mtase and RdRp complex was subjected to further molecular dynamics studies (20 ns). The procedure for Molecular dynamics simulation was the same as in “Molecular Dynamics (MD) Simulations” Section and thermodynamic calculations as “Binding Free Energy Calculations ”Section.

2.9 Assessment of drug likeness

The online software SwissADME was used to compute the physicochemical descriptors as well as predict the pharmacokinetic properties and drug-like nature of the screened compounds compared to that of BG323 and Ribavirin. to (Bultet *et al.*, 2016; Daina, Michielin, & Zoete, 2014). SwissADME utilizes the “Brain Or Intestinal Estimated permeation, (BOILED-Egg)” method which computes the lipophilicity and polarity of small molecules (Daina & Zoete, 2016).

3. Results and Discussion

3.1 Homology Model and Active Binding Site Determination:

Due to the absence of a crystal structure for the Zika NS5 enzyme, a homology model, having a zDope score of -0.76 was generated, and validated using a ramachandran plot. The active site residues were determined for both the MTase and RdRp region (Figure S1-3). The comprehensive set of results are presented in our previous publication (Ramharack & Soliman, 2016). To further validate both the MTase and RdRp, the homology model was superimposed to the newly released crystal structure of the Zika NS5 (Figure 5.2), using Chimera (E F Pettersen *et al.*, 2004).

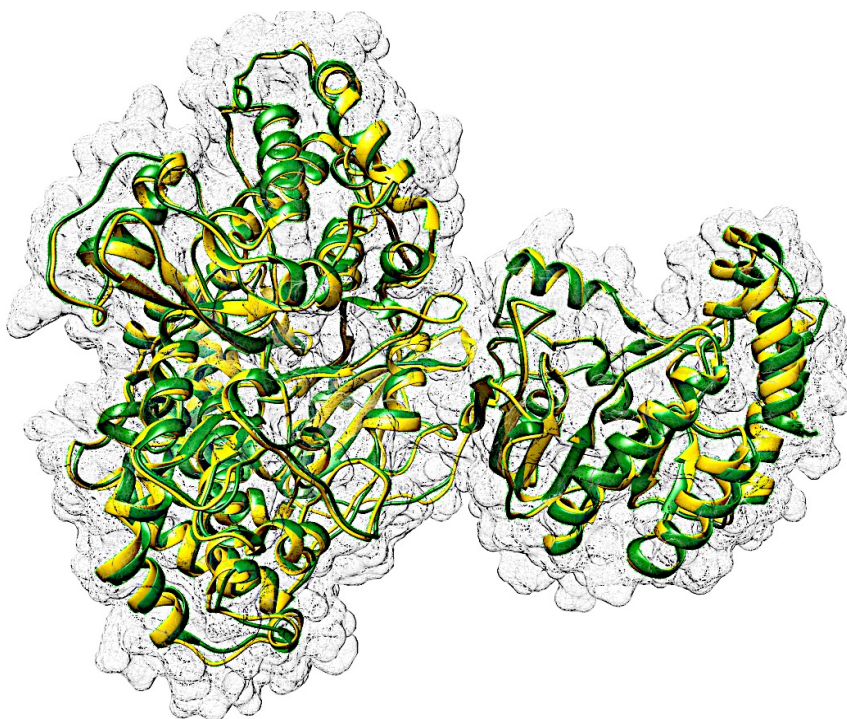


Figure 5.2: Superimposition of homology model (yellow) with the newly released crystal structure (green- PDB code: 5TFR), showing their structural similarity and validating the model's use for subsequent analysis.

3.2 PRED Pharmacophore Model:

In this study, a pharmacophore hypothesis was adopted by utilizing per residue decomposition energy-based approach. The structural features of a protein as well as the chemical characteristics

of a ligand are employed in the construction of a pharmacophore model. To generate the pharmacophore model, a 5ns molecular dynamic simulation was run on complexes (MTase-BG323) and (RdRp-Ribavirin), followed by PRED computed from MM/GBSA calculations. The MM/GBSA approach has proven to be, in principle, accurate in both scoring function and binding free energy results (Genheden & Ryde, 2015; Hayes, 2009). This allows for improved pharmacophore modeling and thus the generation of a concise library of small molecules. The MTase-BG323 complex showed His104 (-2.176 kcal/mol), Glu143 (-1.846 kcal/mol), Thr210 (-1.192 kcal/mol), and Lys176 (-1.061 kcal/mol) to be the highest contributing residues to interact with the ligand. Strong hydrophobic interactions were formed between Glu143 and the benzene ring of BG323, while, energetically favorable residue, Asp140, formed hydrogen bonds with the terminal hydroxyl groups of the ligand (Figure 5.3).

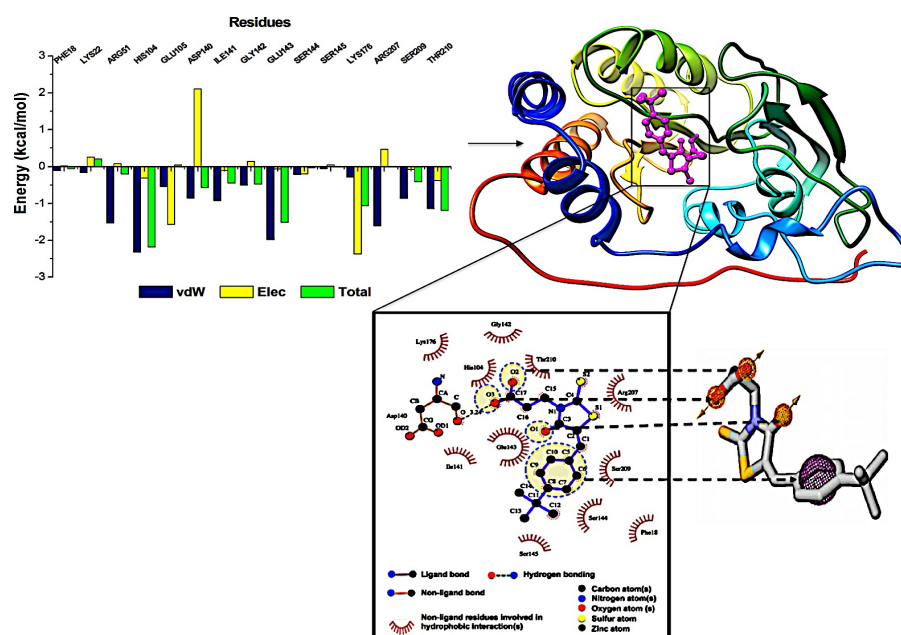


Figure 5.3: The steps taken toward creating the pharmacophore model from the MTase-BG323 complex. The yellow circles spotlight the pharmacophoric moieties that were chosen for the model, based on the highest contributing residues, depicted in the binding affinity graph.

Hydrogen bonds between the RdRp-Ribavirin complex included residues Asn444 (-1.296 kcal/mol) and His460 (-0.956 kcal/mol), while the contributing residues; namely, GLU573 (-1.521 kcal/mol), TRP576 (-1.744 kcal/mol) and Cys577 (-2.202 kcal/mol) were involved in hydrophobic interactions with the ligand. The features from each complex were used as a query on ZINCpharmer (Koes & Carlos, 2012) to create the PRED-based pharmacophore (Figure 5.4). Results revealed 18 hits obtained from the MTase-BG323 pharmacophore and 23 hits from the RdRp-Ribavirin pharmacophore.

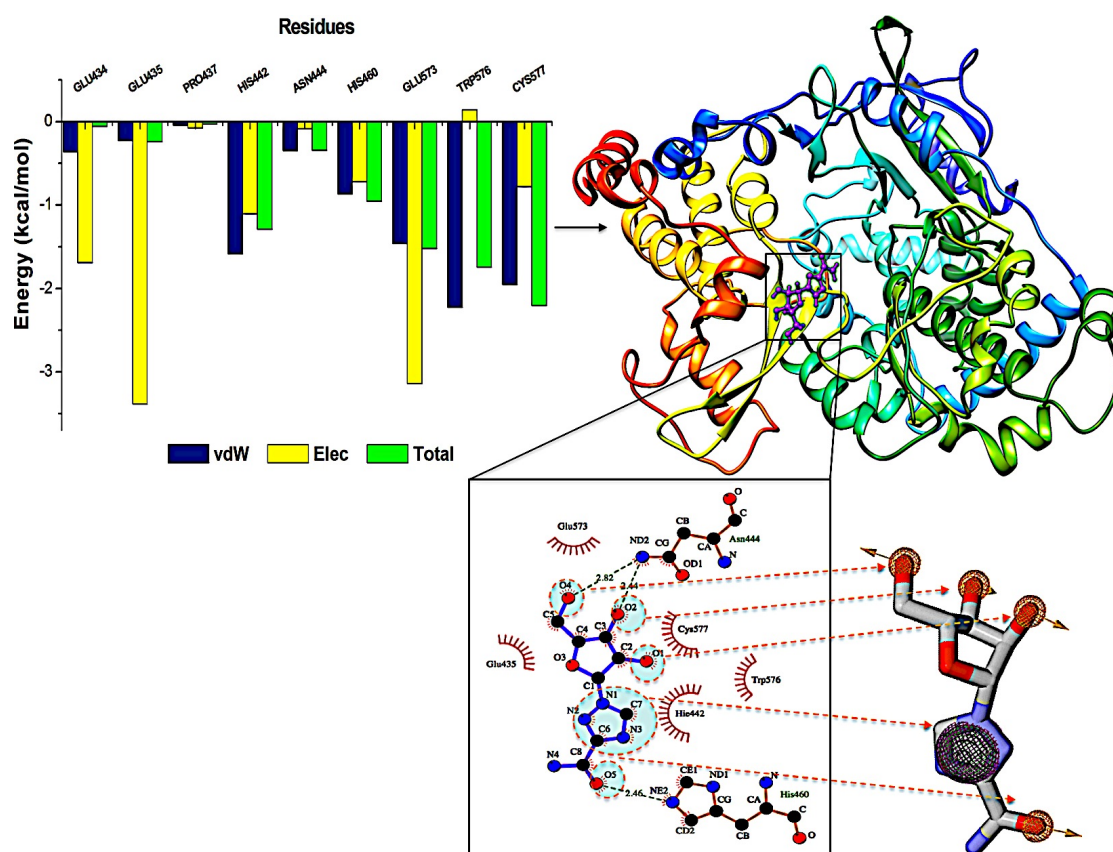
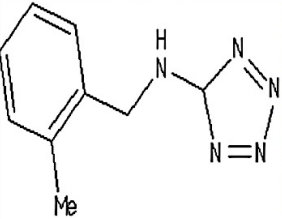
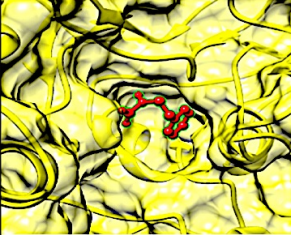
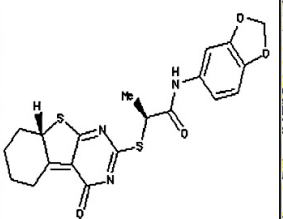
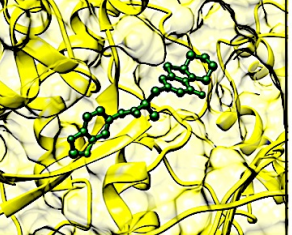
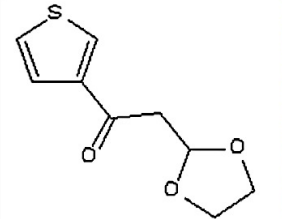
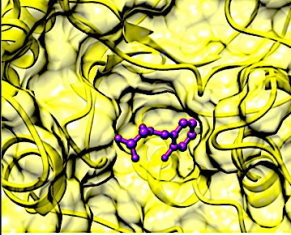
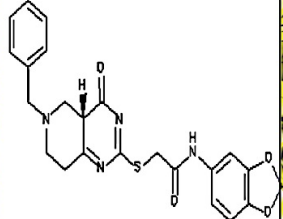
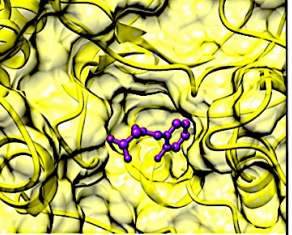
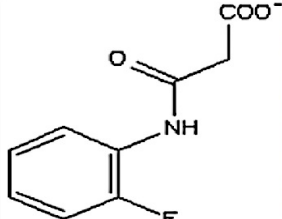
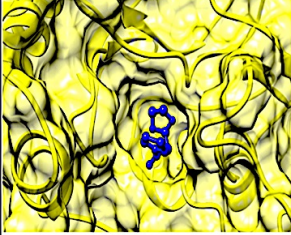
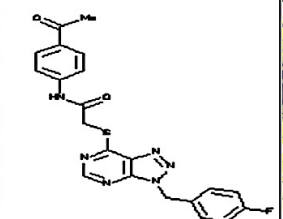
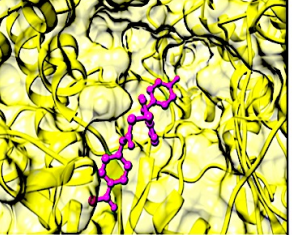


Figure 5.4: RdRp-Ribavirin complex ligplot analysis- creating the pharmacophore model to virtually screen for new RdRp potential lead compounds. The yellow circles spotlight the pharmacophoric moieties that were chosen for the model based on the highest contributing residues, depicted in the binding affinity graph.

3.3 Molecular Docking

To further refine and reduce false positives retrieved from the hit compounds, the hits for each complex were subjected to molecular docking within the active sites of MTase (18 hits) and RdRp (23 hits). This assessed their geometric feasibility at each domain, leading to only three top ranked compounds (Table 2). Based on the interactions and binding affinities of the respective three top ranked compounds to MTase and RdRp, ZINC64717952 and ZINC39563464 were chosen as respective top hits. Each complex was subsequently subjected to molecular dynamic studies to elucidate on the enzyme-ligand interactions of the two potential inhibitors under virtual conditions.

Table 5.2: Representation of the top three compounds bound to MTase and RdRp. The compounds, ZINC64717952 and ZINC39563464 showed the best binding affinity to MTase and RdRp respectively.

MTase		RdRp	
Compound	Docked structure	Compound	Docked structure
 ZINC64717952	 $\Delta G = -6.1$ kcal/mol	 ZINC39563464	 $\Delta G = -9.1$ kcal/mol
 ZINC85652269	 $\Delta G = -5.4$ kcal/mol	 ZINC39588253	 $\Delta G = -7.9$ kcal/mol
 ZINC09304524	 $\Delta G = -5.1$ kcal/mol	 ZINC11758496	 $\Delta G = -7.5$ kcal/mol

3.4 Molecular Dynamic Simulations and Binding Free Energy Analysis

The MTase-ZINC64717952 and RdRp-ZINC39563464 complexes were subjected to a 20 ns MD simulation in order to check the convergence dynamic stability and to analyze the energetics of each complex. The RMSD profiles of the MTase-complex and RdRp-complex indicate that both systems were stable during the simulation (Figure 5.6C and 5.7C).

MTase-ZINC64717952 Complex:

The docked MTase-ZINC64717952 complex showed ionic interactions involving seven residues common to MTase-BG323 (Glu143, Arg207, Lys176, Thr210, Ile141, Asp140 and Gly142). Interestingly, however, a hydrogen bond was noted between the nitrogen of Arg35 and the aromatic ring of ZINC64717952, this was peculiar, as Arg35 was not involved in any ionic interactions of the MTase-BG323 complex. The MTase-ZINC64717952 complex used Asp140 as a hydrogen bond acceptor, whereas, the MTase-BG323 complex depicted hydrophobic interactions between Asp140 and the benzene ring of BG323. These ionic bond deviations between systems may be due to the size of ZINC64717952 in comparison to BG323. ZINC64717952 was significantly reduced in size, containing predominantly the heterocyclic rings from the pharmacophore model. Due to the size of ZINC64717952, the nitrogen of aromatic ring was allowed to form a hydrogen bond with the amine group (Arg35) further into the hydrophobic pocket of MTase. Docking results showed the same binding affinity in both complexes, however, receptor residue stability showed increased fluctuations in the ZINC64717952-MTase complex compared to the experimental complex (Figure 5.5). The overall compactness of the receptor was measured by the radius of gyration (around the C α atoms) and was indicative of greater fluctuations of the MTase-ZINC64717952 complex compared to the experimental complex (Figure S6), verifying the RMSF fluctuations seen in Figure 5.5. Although ZINC64717952 docked in a structurally favorable manner, MM/GBSA analysis showed free

binding energy of the MTase-BG323 complex (-28.70 kcal/mol) to be higher in magnitude than that of MTase-ZINC64717952 (-26.50 kcal/mol).

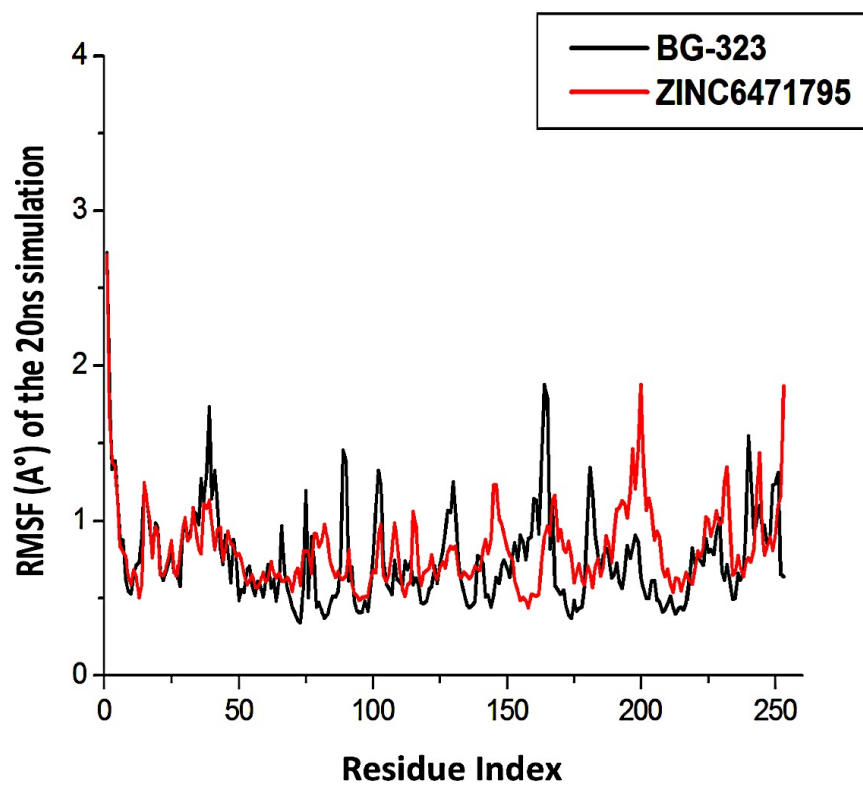
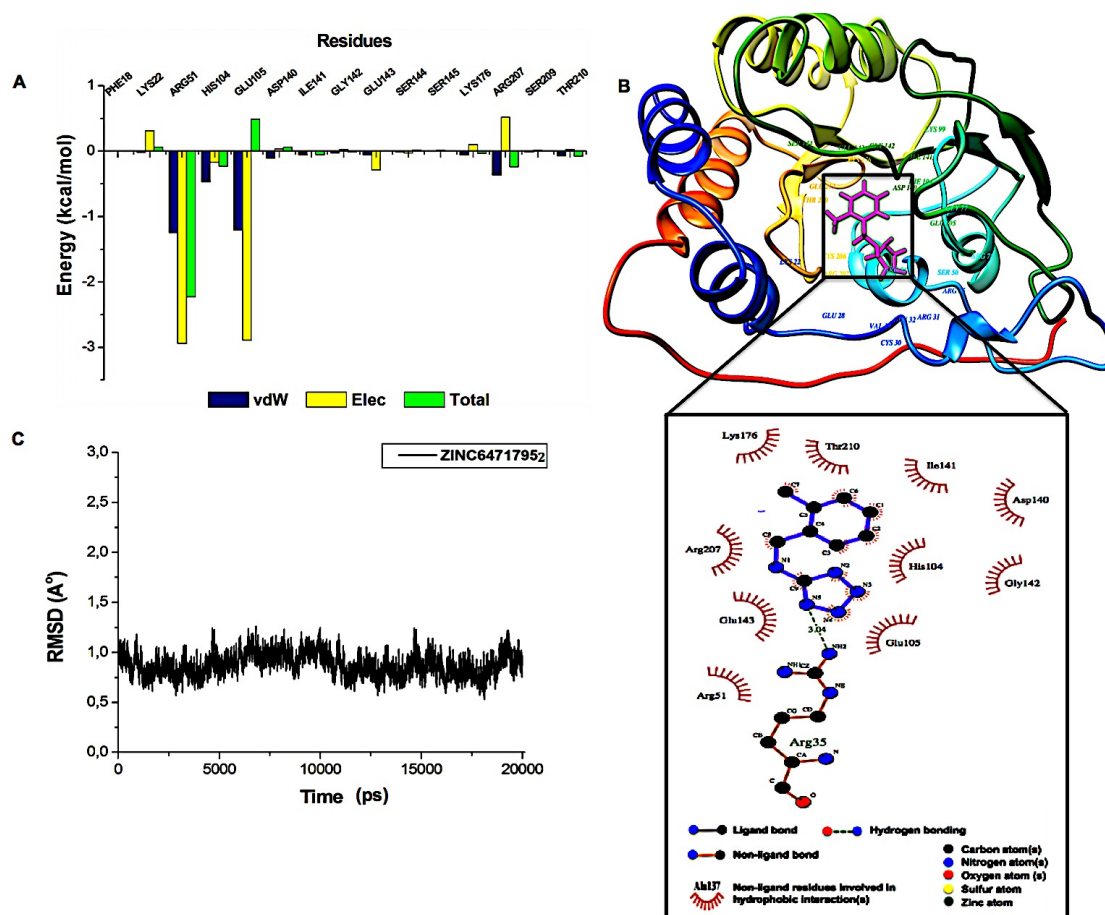


Figure 5.5: The C α root mean square fluctuations (RMSF) of MTase-BG323 and MTase-ZINC64717952 during the molecular dynamic simulation.

The tetrazole aromatic ring in ZINC64717952 contains highly active nitrogen atoms, increasing electronegativity and steric hindrance (Ostrovskii, Trifonov, & Popova, 2012). The Generalized Born (GB) method is used to calculate the molecular electrostatic forces in solvent. Table 3 shows ZINC64717952 to have elevated columbic energy, thus leading to increased gas-phase energy, validating the free energy analysis (Figure 5.6A) (Genheden & Ryde, 2015). This, however, does not rule-out the possibility of ZINC64717952 as a potential inhibitor of the MTase enzyme as the intermolecular forces between the receptor and ligand were favorable. This study

will have to be evaluated *in vitro*, where the further analysis may reveal the inhibitory potential of the compound.



5. 5.6: MTase-ZINC64717952 complex interactions (A) Per-residue decomposition analysis showing Arg51 and Glu105 to have the greatest bond fluctuations (B) Ligplot depiction of hydrophobic and hydrogen bond interactions in the complex which was validated by (C) The time evolution of RMSD of the C-alpha atom backbone of the MTase-ZINC64717952 complex.

Table 5.3: The comparison of MTase's binding affinity with BG323 and ZINC64717952.

Energy Components (kcal/mol)					
Compound	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
BG323	-33.32 \pm	-11.84 \pm	-49.16 \pm	20.83 \pm 1.80	-28.33 \pm
	1.82	1.68	1.86		1.87
ZINC64717952	-35.77 \pm	-10.47 \pm	-46.24 \pm	19.74 \pm 2.14	-26.50 \pm
	2.66	2.68	3.97		3.14

RdRp-ZINC39563464 Complex:

The docked RdRp-ZINC39563464 complex showed ZINC39563464 to interact with nitrogen atoms of two residues; Asn444 and His460. Notably, the nitrogen atoms from the same residues form hydrogen bonds with the terminal oxygen of the Ribavirin, showing consistent residue interactions of the experimental ligand and ZINC39563464. These hydrogen interactions are formed from non-covalent bonding of the hydrogen donor (Asn444 and His460) with the acceptors (oxygen and nitrogen) of the ligand. This articulates the directionality and specificity of the active site's β -strand recognition of both Ribavirin and ZINC39563464. The complex exhibiting a relatively stable RMSD profile during the simulation further validated this (Figure 5.7C). The pharmacophoric hot spot residue, His442, formed hydrophobic bonds with the aromatic rings of both Ribavirin and ZINC39563464. It is noteworthy that four other hydrophobic-interacting residues; Cys577, Tryp576, Glu573 and Glu435 were common to both ligands, thus stabilizing both energetically favorable ligands in the available hydrophobic pocket. Table 4 depicts the analysis of binding free energy by the use of MM/GBSA of the RdRp-ZINC39563464

complex was used to support the docking results.

Table 5.4: The comparison of RdRp's binding affinity with Ribavirin and ZINC39563464.

Energy Components (kcal/mol)					
Compound	ΔE_{vdw}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
Ribavirin	-23.20 ±	-40.92 ±	-64.12 ±	47.59 ±	-16.53 ±
	3.13	13.03	13.34	9.93	4.84
ZINC39563464	-38.17 ±	-17.32 ±	-55.49 ±	30.45 ±	-25.04 ±
	5.39	5.99	8.85	4.85	5.35

The predicted binding free energy for the complex was -25.04 kcal/mol, which is considerably higher in magnitude than that achieved by the RdRp-Ribavirin complex (-16.53kcal/mol), thus confirming the docking results and indicating a stronger binding of ZINC39563464 to RdRp compared to the experimental ligand (Figure 5.7A and 5.7B). The relatively large size of the ligand could explain the increased number of residues encompassing apparent hydrophobic interactions with ZINC39563464, and could substantiate the exhibition of stable RdRp residues by RMSF profiling (Figure 5.8). As an additional check, the radius of gyration (RoG of the Cα atoms) was compared in both simulations to provide a measure of overall compactness of the protein (Figure S6). The fluctuations of RoG stayed with 1 Å in both simulations indicative of a stable protein complex with both experimental and screened compound.

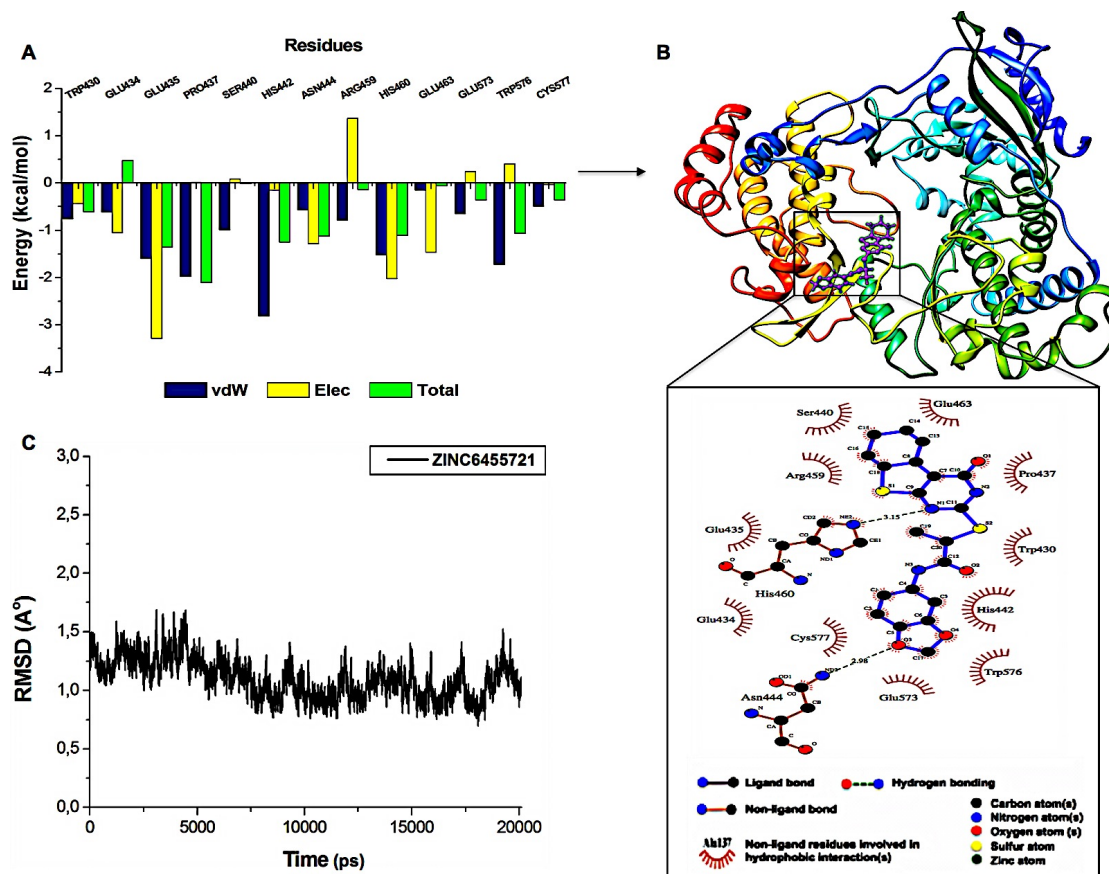


Figure 5.7: RdRp-ZINC39563464 complex interactions (A) Per-residue decomposition analysis showing Arg459 and Glu435 to have the greatest bond fluctuations, (B) Ligplot depiction of hydrophobic and hydrogen bond interactions in the complex which was validated by (C) The time evolution of RMSD of the C-alpha atom backbone of the MTase-ZINC39563464 complex.

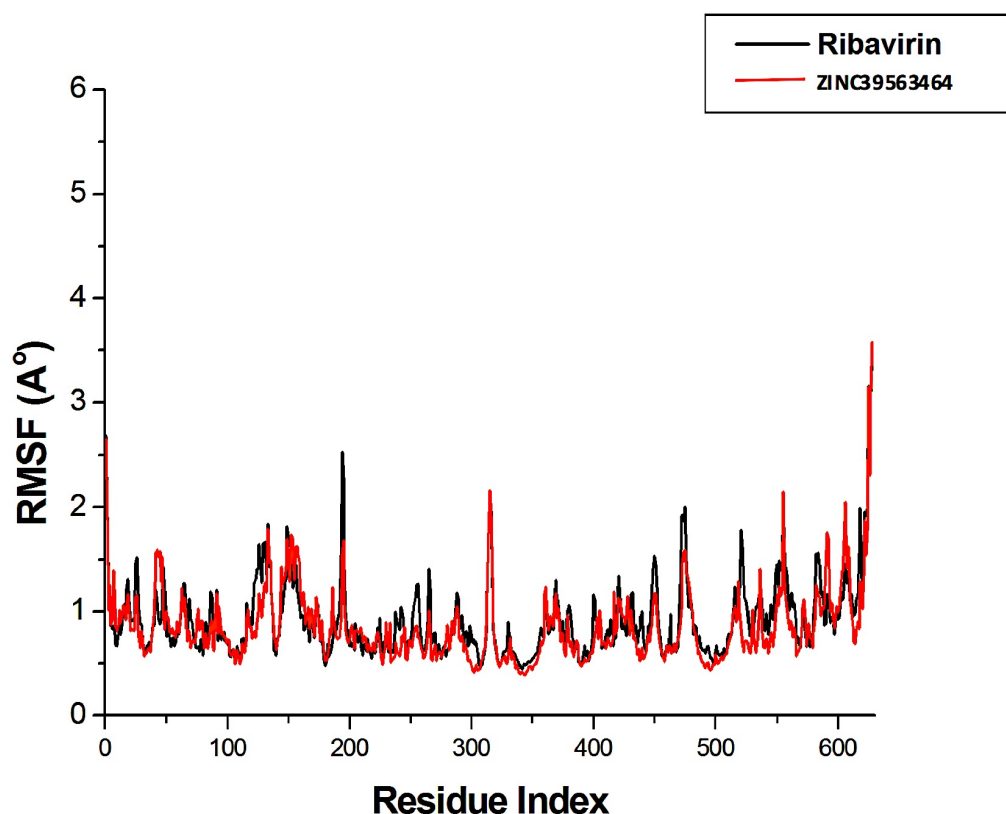
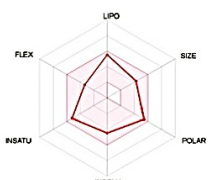

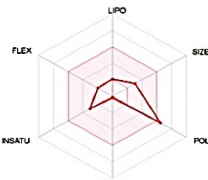
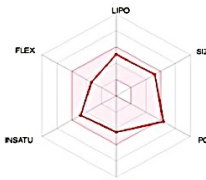


Figure 5.8: The stable C α root mean square fluctuations (RMSF) of RdRp-Ribavirin and RdRp-ZINC39563464 during the molecular dynamic simulation.

3.4 Assessment of drug likeness

Ribavirin has a plethora of side effects including thrombocytopenia, myalgia, leucopenia and cognitive impairment (Hinton *et al.*, 2016; Kryger, Wohl, Smith, & Zelikin, 2013; Munir *et al.*, 2010). This proves to be a challenge when trying to inhibit a virus that already causing these symptoms. BG323 is a new compound that has been proven to have potent effects on *flavivirus* NS5 proteins, however, the compound is unable to pass the blood-brain barrier, making it difficult to act on ZIKV-targeted neuronal cells (Miner & Diamond, 2016; Tang *et al.*, 2016). The possible pro-drugs of ribavirin and BG323, being, ZINC39563464 and ZINC64717952 respectively, can be described as potential lead compounds after assessment through SwissADME (Table 5) (Bultet *et al.*, 2016).

Table 5.5: The comparison of drug likeness of the screened compounds compared to that of the experimental drugs against ZIKV.

	BG323	ZINC6471795	Ribavirin	ZINC39563464
Molecular formula	C17H19N03S2	C9H11N5	C8H12N4O5	C20H19N3O4S2
Molecular weight g/mol	349.47	189.22	244.20	429.51
Lipophilicity (iLOGP)	2.79	2.47	0.22	2.86
Water soluble	Moderately soluble	Soluble	Soluble	Moderately Soluble
GIT absorption	High	High	Low	Low
BBB permeability	Not permeable	Semi-permeable	Not permeable	Semi-Permeable
Bioavailability score	0.56	0.55	0.55	0.55
Synthetic accessibility	3.39	2.93	3.89	4.89
Druglikeness (Lipinski)	Yes	Yes	Yes	Yes
"Boiled-Egg" method summary				

4. Conclusion

ZIKV is a rapidly evolving virus that has had detrimental long-term effects over a very short period of time. This study proposes two new compounds that have shown promising physicochemical properties and strong interactions with ZIKV MTase and RdRp, thus validating the PRED model as an effective strategy to enhance typical virtual screening methods for the rapid identification of potential lead compounds as inhibitors against pathogenic biological targets such as ZIKV. This strategic *in silico* technique will serve as a beneficial tool to enhance drug discovery and decrease excessive wastage of financial and experimental resources by synthesizing large numbers of compounds that may not be beneficial in the inhibition of target enzymes. The lead compounds, ZINC64717952 and ZINC39563464, have shown substantial stability in complex with the target enzymes and thus further experimental analysis is necessary for efficacy and toxicity validation.

Conflicts of interest

Authors declare no potential conflicts of interest.

References

- Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Philip, E., Burkhardt, K., ... Westbrook, J. D. (2002). The Protein Data Bank. *Biological Crystallography*, 58, 899–907.
- Brecher, M., Chen, H., Liu, B., Banavali, N. K., Jones, S. A., Zhang, J., ... Li, H. (2015). Novel broad spectrum inhibitors targeting the flavivirus methyltransferase. *PLoS ONE*, 10(6), 1–15.
- Brito, C. (2016). Zika Virus: A New Chapter in the History of Medicine. *The Lancet Neurology*, 28(6), 679–680.
- Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, 41, 349–357.
- Bultet, L. A., Aguilar-Rodriguez, J., Ahrens, C. H., Ahrn??, E. L., Ai, N., Aimo, L., ... Zollinger, A. (2016). The SIB Swiss Institute of bioinformatics' resources: Focus on curated databases. *Nucleic Acids Research*, 44(D1), D27–D37.
- Campos, G. S., Bandeira, A. C., & Sardi, S. I. (2015). Zika virus outbreak, Bahia, Brazil. *Emerging Infectious Diseases*, 21(10), 1885–1886.
- Cele, F. N., Muthusamy, R., & Soliman, M. E. S. (2016). Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: a study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. *Drug Design, Development and Therapy*, 10, 1365–1377.

- Chen, B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., ... Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66(1), 12–21.
- Chibueze, E. C., Tirado, V., & Olukunmi, O. (2016). Zika virus infection in pregnancy : a systematic review of disease course and complications. *Bulletin of the World Health Organization*, (June), 1–35.
- Daina, A., Michielin, O., & Zoete, V. (2014). ILOGP: A simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. *Journal of Chemical Information and Modeling*, 54(12), 3284–3301.
- Daina, A., & Zoete, V. (2016). A BOILED-Egg To Predict Gastrointestinal Absorption and Brain Penetration of Small Molecules. *ChemMedChem*, 1117–1121.
- Egloff, M. P., Benarroch, D., Selisko, B., Romette, J. L., & Canard, B. (2002). An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: Crystal structure and functional characterization. *European Molecular Biology Organization Journal*, 21(11), 2757–2768.
- Ekins, S., Mietchen, D., Coffee, M., Stratton, T., Freundlich, J., Freitas-Junior, L., ... Andrade, C. (2016). Open drug discovery for the Zika virus. *FL1000Research*, 5(150), 1–14.
- Eswar, N., Webb, B., Marti-Renom, M. a, Madhusudhan, M. S., Eramian, D., Shen, M. Y., ... Sali, A. (2006). Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics*.
- Faye, O., Freire, C. C. M., Iamarino, A., Faye, O., de Oliveira, J. V. C., Diallo, M., ... Sall, A. A. (2014). Molecular Evolution of Zika Virus during Its Emergence in the 20th Century. *PLoS*

Neglected Tropical Diseases, 8(1), 1–10.

- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols*, 11(5), 905–919.
- Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, 10(5), 449–61.
- Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., ... Weaver, S. C. (2012). Genetic characterization of zika virus strains: Geographic expansion of the asian lineage. *PLoS Neglected Tropical Diseases*, 6(2), 1–7.
- Hayes, E. B. (2009). Zika virus outside Africa. *Emerging Infectious Diseases*, 15(9), 1347–1350.
- Hernandez, M., Ghersi, D., & Sanchez, R. (2009). SITEHOUND-web: A server for ligand binding site identification in protein structures. *Nucleic Acids Research*, 37(SUPPL. 2), 413–416.
- Hinton, T. M., Zuwala, K., Deffrasnes, C., Todd, S., Shi, S., Marsh, G. A., ... Zelikin, A. N. (2016). Polyanionic Macromolecular Prodrugs of Ribavirin: Antiviral Agents with a Broad Spectrum of Activity. *Advanced Healthcare Materials*, 5(5), 534–540.
- Idrus, S., Tambunan, U., & Zubaidi, A. A. (2012). Designing cyclopentapeptide inhibitor as potential antiviral drug for dengue virus ns5 methyltransferase. *Bioinformation*, 8(8), 348–352.
- Irwin, J. J., & Shoichet, B. K. (2005). ZINC - A free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*, 45(1), 177–182.
- Koes, D. R. C., & Carlos, J. (2012). ZINCPharmer: Pharmacophore search of the ZINC database.

Nucleic Acids Research, 40(W1), 409–414.

Kroemer, R. T. (2007). Structure-based drug design: docking and scoring. *Current Protein & Peptide Science*, 8(4), 312–328

Kryger, M. B. L., Wohl, B. M., Smith, A. a a, & Zelikin, A. N. (2013). Macromolecular prodrugs of ribavirin combat side effects and toxicity with no loss of activity of the drug. *Chemical Communications (Cambridge, England)*, 49(26), 2643–5.

Kumalo, H. M., & Soliman, M. E. (2015). Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery: A Case Study on the Identification of Novel β -Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents. *Cellular and Molecular Bioengineering*, 9(1), 175–189.

Kusumaningrum, S., Budianto, E., Kosela, S., Sumaryono, W., & Juniarti, F. (2014). The molecular docking of 1,4-naphthoquinone derivatives as inhibitors of Polo-like kinase 1 using Molegro Virtual Docker. *Journal of Applied Pharmaceutical Science*, 4(11), 47–53.

Leyssen, P., De Clercq, E., & Neyts, J. (2000). Perspectives for the treatment of infections with Flaviviridae. *Clinical Microbiology Reviews*, 13(1), 67–82.

Lim, S., & Shi, P. (2013). West Nile virus drug discovery. *Viruses*, 5(12), 2977–3006.

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2012). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development setting. *Advanced Drug Delivery Reviews*, 64, 4–17.

Lissauer, D., Smit, E., & Kilby, M. D. (2016). Zika virus and pregnancy. *An International Journal of Obstetrics and Gynaecology*, 1258–1263.

Madden, T. L., Tatusov, R. L., & Zhang, J. (1996). Application of Network BLAST Server. *Methods in Enzymology*, 266, 131–141.

- Mahfuz, M., Khan, A., Mahmud, H., Hasan, M., Parvin, A., Rahman, N., & Rahman, S. M. B. (2014). Indian Journal of Pharmaceutical and Biological Research (IJPBR) In Silico Modeling and Immunoinformatics Probing Disclose the Epitope Based Peptide Vaccine Against Zika Virus Envelope Glycoprotein. *Indian Journal of Pharmaceutical Biological Research*, 2(4), 44–57.
- Mahfuz, M., Khan, A., Mahmud, H., Hasan, M., Parvin, A., Rahman, N., & Rahman, S. M. B. (2015). Design and Prediction of Potential RNAi (siRNA) Molecules for 3' UTR PTGS of different strains of Zika Virus: A Computational Approach. *Nature and Science*, 1(2), 37–50.
- Millichap, J. G. (2016). Zika Virus Infection and Microcephaly. *Pediatric Neurology Briefs*, 30(1), 8.
- Miner, J. J., & Diamond, M. S. (2016). Understanding how zika virus enters and infects neural target cells. *Cell Stem Cell*.
- Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., ... Avšič Županc, T. (2016). Zika Virus Associated with Microcephaly. *New England Journal of Medicine*, 374(10), 951–958.
- Morris, G., & Huey, R. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791.
- Mumtaz, N., van Kampen, J. J. A., Reusken, C. B. E. M., Boucher, C. A. B., & Koopmans, M. P. G. (2016). Zika Virus: Where Is the Treatment? *Current Treatment Options in Infectious Diseases*, 8(3), 208–211.
- Munir, S., Saleem, S., Idrees, M., Tariq, A., Butt, S., Rauff, B., ... Awan, Z. (2010). Hepatitis C Treatment: current and future perspectives. *Virology Journal*, 7, 296.

- Nair, P. C., & Miners, J. O. (2014). Molecular dynamics simulations: from structure function relationships to drug discovery. *In Silico Pharmacology*, 2(4), 1–4.
- Ostrovskii, V. A., Trifonov, R. E., & Popova, E. A. (2012). Medicinal chemistry of tetrazoles. *Russian Chemical Bulletin*, 61(4), 768–780.
- Panchaud, A., Stojanov, M., Ammerdorffer, A., & Vouga, M. (2016). Emerging Role of Zika Virus in Adverse Fetal and Neonatal Outcomes. *Clinical Microbiological Reviews*, 29(3), 659–694.
- Papageorgiou, L., Loukatou, S., Koumandou, V. L., Makałowski, W., Megalooikonomou, V., Vlachakis, D., & Kossida, S. (2014). Structural models for the design of novel antiviral agents against Greek Goat Encephalitis. *PeerJ*, 2, 1–18.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, 25(13), 1605–1612.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612.
- Ramharack, P., & Soliman, M. E. S. (2016). Zika virus drug targets: a missing link in drug design and discovery – a route map to fill the gap. *RSC Advances*, 6(73), 68719–68731.
- Roa, M. (2016). Zika virus outbreak: Reproductive health and rights in Latin America. *The Lancet*, 387(10021), 843.
- Sanner, M. F., Olson, J., & Spehner, J. C. (1996). Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*, 38(3), 305–320.
- Shanmugam, A., Velmurugan, D., & Gromiha, M. M. (2016). Identification of dengue viral

- RNA-dependent RNA polymerase inhibitor using computational fragment-based approaches and molecular dynamics study. *Journal of Biomolecular Structure & Dynamics*, 34, 1512-1532.
- Shapshak, P., Sinnott, J. T., Somboonwit, C., & Kuhn, J. H. (2016). Global virology I-identifying and investigating viral diseases. *Global Virology I-Identifying and Investigating Viral Diseases*, (Chapter 18), 477–500. 3
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7(1), 539.
- Singh, R. K., Dhama, K., Malik, Y. S., Ramakrishnan, M. A., Karthik, K., Tiwari, R., Joshi, S. (2016). Zika Virus - Emergence, evolution, pathology, diagnosis and control: current global scenario and future perspectives - A comprehensive review. *The Veterinary Quarterly*, 2176(July), 1–43.
- Soliman, M. E. S. (2013). A hybrid structure/pharmacophore-based virtual screening approach to design potential leads: A computer-aided design of South African HIV-1 subtype C protease inhibitors. *Drug Development Research*, 74(5), 283–295.
- Sweeney, N. L., Hanson, A. M., Mukherjee, S., Ndjomou, J., Geiss, B. J., Steel, J. J., ... Frick, D. N. (2015). Benzothiazole and Pyrrolone Flavivirus Inhibitors Targeting the Viral Helicase. *ACS Infectious Diseases*, 1(3), 140–148.
- Tambunan, U. S. F., Zahroh, H., Utomo, B. B., & Parikesit, A. A. (2014). Screening of commercial cyclic peptide as inhibitor NS5 methyltransferase of Dengue virus through Molecular Docking and Molecular Dynamics Simulation. *Bioinformation*, 10(1), 23–27.
- Tang, H., Hammack, C., Ogden, S. C., Wen, Z., Qian, X., Li, Y., ... Ming, G. L. (2016). Zika

- virus infects human cortical neural progenitors and attenuates their growth. *Cell Stem Cell*, 18(5), 587–590.
- Troncoso, A. (2016). Zika threatens to become a huge worldwide pandemic. *Asian Pacific Journal of Tropical Biomedicine*, 6(6), 520–527.
- Trott, O., & Olson, A. J. (2010). AutoDock Vina. *J. Comput. Chem.*, 31, 445–461.
- Turmel, M. J., Hubert, M. J. P., Maquart, Y. M. V, Guillou-Guillemette, M., & Leparac-Goff, I. (2016). Late sexual transmission of Zika virus related to. *The Lancet*, 6736(16), 30775.
- Wang, J., Wang, W., Kollman, P. A., & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, 25(2), 247–260.
- WHO. (2016). Zika virus, Microcephaly and Guillain-Barré syndrome, 1–12.
- Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C. C., Pettersen, E. F., ... Ferrin, T. E. (2012). UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *Journal of Structural Biology*, 179(3), 269–278.
- Zmurko, J., Marques, R. E., Schols, D., Verbeken, E., Kaptein, S. J. F., & Neyts, J. (2016). The Viral Polymerase Inhibitor 7-Deaza-2-C-Methyladenosine Is a Potent Inhibitor of In Vitro Zika Virus Replication and Delays Disease Progression in a Robust Mouse Infection Model. *PLoS Neglected Tropical Diseases*, 10(5), 1–15.
- Zou, X. W., Liu, Y. C., Hsu, N. S., Huang, C. J., Lyu, S. Y., Chan, H. C., ... Li, T. L. (2014). Structure and mechanism of a nonhaem-iron SAM-dependent C-methyltransferase and its engineering to a hydratase and an O-methyltransferase. *Acta Crystallographica Section D: Biological Crystallography*, 70(6), 1549–1560.

CHAPTER 6

Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery

Pritika Ramharack^A, Sofiat Oguntade^A Mahmoud E. S. Soliman^{A*}

^AMolecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

1. Dean and Head of School of Health Sciences, Full Professor: Pharmaceutical Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.
2. Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt.
3. College of Pharmacy and Pharmaceutical Sciences, Florida Agricultural and Mechanical University, FAMU, Tallahassee, Florida 32307, USA.

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Abstract

The Zika virus has emerged as a pathogen of major health concern. The rapid spread of the virus has led to an uproar in the medical domain as scientists frantically race to develop effective vaccines and small molecules to inhibit the virus. In the past year, there has been a flood of Zika knowledge published including its characteristics, transmission routes and its role in disease conditions such as Microcephaly and Gullian-Barre syndrome. Targeted therapy against specific viral maturation proteins is necessary in halting the replication of the virus in the human host, thus decreasing host-host transmission. This prompted us to investigate the structural properties of the Zika NS3 Helicase when bound to ATP-competitive inhibitor, NITD008. In this study, comparative molecular dynamic simulations were employed for Apo and bound protein to demonstrate the molecular mechanism of the Helicase. Results clearly revealed that NITD008-binding caused significant residue fluctuations at the P-loop compared to the rigid nature of the Apo conformation. The NITD008-helicase complex also revealed residues 339-348 to transition from a $_{310}$ -Helix to a stable α -helix. These protein fluctuations were verified by investigation of dynamic cross correlation and principal component analysis. The fundamental dynamic analyses presented in this report is crucial in understanding Zika NS3 Helicase function, thereby giving insights toward an inhibition mechanism. The information reported on the binding mode at the ATPase active site may also assist in designing of effective inhibitors against this detrimental viral target.

Keywords:

Zika NS3 Helicase, ATPase active site, Molecular dynamic simulations, P-Loop Flexibility,

Introduction

The re-emerging Zika virus (ZIKV) has evolved into a catastrophic epidemic over the past year, with scientific community announcing that the long-term effects associated with the virus will have to be dealt with in the decades to follow ¹. The virus was declared an international public health emergency by the World Health Organization ², based on growing evidence of the virus being linked with congenital neurological diseases such as Guillain-Barre, cranial nerve dysfunction and Microcephaly ^{3,4}. The ZIKV made its devastating re-appearance in Brazil and has now spread on a global scale, with an estimated 75 countries with reported mosquito-borne ZIKV transmission as of December 2016 ⁵.

Zika virus is an arthropod-borne *flavivirus* initially discovered in the Zika forest area of Uganda in 1947 ⁶. Of the *flavivirus* genera, ZIKV is most closely related to the Spodweni virus from the Spodweni group; however, ZIKV shares structural similarities with other *flaviviruses*, including Dengue virus and West Nile virus ⁷. The ZIKV genome is made up of structural proteins, being the capsid, precursor membrane and envelope form the viral particle and seven non-structural proteins, being NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5, which participate in the replication of the RNA genome, virion assembly and invasion of the innate immune system ⁸⁻¹⁰. In our previous review, we explicated on the key viral target proteins, including the multifunctional viral replication NS3 helicase protein¹¹. The ZIKV helicase comes from the superfamily helicases, SF2 ¹², with the inhibition of either one of the binding sites, the RNA-binding groove or the ATP-binding site (Figure 6.1), leading to the virus becoming incapable of sufficient maturation and replication. The structural characteristics of the ZIKV NS3 protein includes three domains: domain I (residues 182-327), domain II (residues 328-480) and domain

III (residues 481-617), as well as a P-Loop (residues 196-203) which is located at the ATP-binding site of domain I ^{12,13}.

The co-crystallization of MnATP⁻² and RNA with ZIKV helicase, reported by Tian *et al* (2016) and Cao *et al* (2016), have paved the way to understanding the mechanism by which these substrates bind to the enzyme, initiating viral RNA replication ^{14,15}. Despite the flood of integrated knowledge on ZIKV over the past year, the molecular and structural mechanism for helicase inhibition is yet to be established ¹².

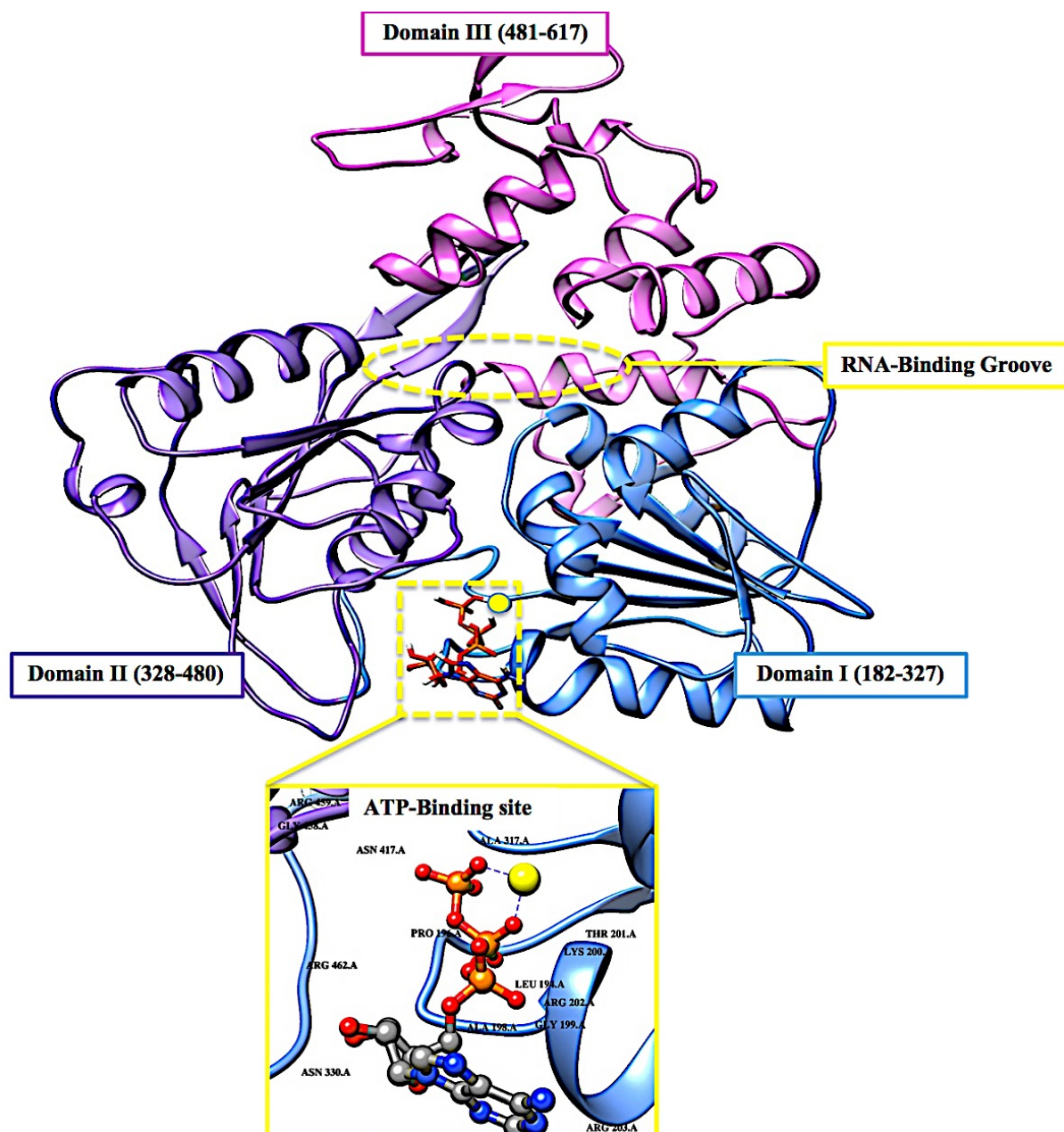


Figure 6.1: Cartoon and surface representation of the three domains of the ZIKV helicase and the two active-binding regions (yellow) that form profound hydrophobic cavities in the electrostatic surface area, allowing ATP and ssRNA to bind.

Another battle being fought by researchers is the discovery of new modes of transmission of the virus, from initially being transmitted from vector to host, to now being inclusive of blood transfers from host to host as well as secondary sexual transmission^{16–18}. This has allowed for

rapid diffusion of the virus between continents. In the plethora of strategic characteristics of the virus, its ability to target neuronal cells has been one of the most problematic tasks that pharmaceutical chemists have had to overcome^{19–24}. The design of ZIKV inhibitors will not only need to be target-specific, effective and have minimal toxicity, but it will also have to pass through the blood-brain-barrier²⁵.

Although there are currently vaccine clinical trials under way, there are still no FDA approved small molecule inhibitors against the virus^{26–30}. This may be due to a number of reasons including time-consuming experimental testing of large libraries of compounds or minimal literature available on the functionality of the virus in host cells. These possible barriers have prompted us to utilize computational drug design tools, such as molecular dynamic (MD) simulations to explore the conformational landscape of this biological system's ATP-binding region. The crystallographic structures have revealed evidence of residue mobility, including the rotation of motor domains, however, the precise structural characteristics of the helicase upon small molecule binding, is yet to be determined^{12,31–36}.

In this study we investigate the conformational changes at the ATP-binding region after a 130 MD simulation of the free enzyme state as well as a NITD008-bound complex³⁷. This study will be critical in understanding how the ZIKV NS3 helicase functions structurally, thus aiding in the design of effective, target-specific inhibitors.

Computational Methods

1.2 System Preparation

The ZIKV NS3 helicase in complex with ATP and a magnesium ion (PDB code: 5GJC)¹⁴ was obtained from RSCB Protein Data Bank³⁸. The 3-D structure of the experimental ZIKV inhibitor, NITD008, was obtained from PubChem³⁹ and prepared on Molegro Molecular Viewer (MMV)⁴⁰. In the ZIKV crystal structure of the ATP-bound helicase, residues A247-S253 were absent, thus the free enzyme (PDB code: 5JMT)¹³ was utilized in the docking of NITD008. Deng *et al* (2016) reported conclusive *in vivo* evidence of the inhibition of ZIKV by NITD008. The compound is classified as an adenosine nucleoside analog that competitively inhibits ATP, thus sharing an active site³⁷.

1.3 Molecular Docking

Molecular docking is a conventional method in computational chemistry which is utilized in the prediction optimized geometric conformations of a ligand within an appropriate binding site⁴¹. The Molecular docking software utilized included Raccoon⁴², Autodock Graphical user interface supplied by MGL tools⁴³ and AutoDockVina⁴⁴ with default docking parameters. Prior to docking, Gasteiger charges were added to NITD008 and the non-polar hydrogen atoms were merged to carbon atoms. Water molecules were removed and polar hydrogen was added to the crystal structure of the NS3 helicase. NITD008 was then docked into the ATPase binding pocket of the NS3 helicase (by defining the grid box with spacing of 1 Å and size of 32 × 26 × 30 pointing in x, y and z directions). Due to the lack of experimental data describing ZIKV approved inhibitors, validation of molecular docking based on the lowest energy pose becomes unreliable⁴⁵. To overcome any experimental bias, the five best conformational poses, based on binding affinities (kcal/mol), were subjected to MD simulations.

1.4 Molecular Dynamic (MD) Simulations

Molecular dynamic (MD) simulations provide a robust tool to explore the physical movements of atoms and molecules, thus providing insights on the dynamical evolution of biological systems. The MD simulation was performed using the GPU version of the PMEMD engine provided with the AMBER package, FF14SB variant of the AMBER force field ⁴⁶ was used to describe the protein.

ANTECHAMBER was used to generate atomic partial charges for the ligand by utilizing the Restrained Electrostatic Potential (RESP) and the General amber Force Field (GAFF) procedures. The Leap module of AMBER 14 allowed for addition of hydrogen atoms, as well as Na⁺ and Cl⁻ counter ions for neutralization to both the Apo- and Bound system.

Both systems were then suspended implicitly within an orthorhombic box of TIP3P water molecules such that all atoms were within 10Å of any box edge.

An initial minimization of 2000 steps was carried out with an applied restraint potential of 500 kcal/mol Å² for both solutes, were performed for 1000 steps using a steepest descent method followed by a 1000 steps of conjugate gradients. An additional full minimization of 1000 steps was further carried out by conjugate gradient algorithm without restrain.

A gradual heating MD simulation from 0K to 300K was executed for 50 ps, such that the system maintained a fixed number of atoms and fixed volume, i.e., a canonical ensemble (NVT). The solutes within the system are imposed with a potential harmonic restraint of 10kcal/mol Å² and collision frequency of 1.0 ps⁻¹. Following heating, an equilibration estimating 500ps of the each system was conducted; the operating temperature was kept constant at 300 K. Additional features such as a number of atoms and pressure where also kept constant mimicking an isobaric-isothermal ensemble (NPT). The systems pressure was maintained at 1 bar using the Berendsen

barostat.

The total time for the MD simulation conducted was 130 ns. In each simulation the SHAKE algorithm was employed to constrict the bonds of hydrogen atoms. The step size of each simulation was 2 fs and an SPFP precision model was used. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of 1 bar maintained by the Berendsen barostat, a pressure-coupling constant of 2 ps, a temperature of 300K and Langevin thermostat with collision frequency of 1.0 ps^{-1} .

1.5 Post-Dynamic Analysis

The coordinates of the free enzyme and NITD008 complex were each saved every 1 ps and the trajectories were analyzed every 1 ps using PTRAJ, followed by analysis of RMSD, RMSF and Radius of Gyration using the CPPTRAJ module employed in AMBER 14 suit.

1.5.1 Binding Free Energy Calculations

Binding free energy calculations is an important end point method that may elucidate on the mechanism of binding between a ligand and enzyme, including both enthalpic and entropic contributions ⁴⁷. To estimate the binding affinity of the docked systems, the free binding energy was calculated using the Molecular Mechanics/GB Surface Area method (MM/GBSA) ⁴⁸. Binding free energy was averaged over 15000 snapshots extracted from the 130 ns trajectory. The free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

$$\Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (4)$$

$$G_{SA} = \gamma SASA \quad (5)$$

The term E_{gas} denotes the gas-phase energy, which consist of the internal energy E_{int} ; Coulomb energy E_{ele} and the van der Waals energies E_{vdw} . The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol} , was estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G . The non-polar solvation energy, G_{SA} , was determined from the solvent accessible surface area (SASA), using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB} , contribution was estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature respectively.

To obtain the contribution of each residue to the total binding free energy profile at the ATPase site, per-residue free energy decomposition was carried out at the atomic level for imperative residues using the MM/GBSA method in AMBER 14 suit.

The system displaying the most favorable binding interaction and energy contributions were subjected to further analysis.

1.5.2 Dynamic Cross-correlation Analysis (DCC)

Dynamic cross correlation is a widespread method in MD simulations in which the correlation coefficients of motions between atoms of a protein may be quantified ⁴⁹. The dynamic cross correlation between the residue-based fluctuations during simulation was calculated using the CPPTRAJ module incorporated in AMBER 14. The formula used to describe dynamic cross correlation is given below:

$$C_y = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{(\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle)^{\frac{1}{2}}}$$

The cross-correlation coefficient (C_{ij}) varies within a range of -1 to $+1$ of which the upper and lower limits correspond to a fully correlated and anti-correlated motion during the simulation process. Where, i and j stands for i^{th} and j^{th} residue respectively and Δr_i or Δr_j represents displacement vectors correspond to i^{th} and j^{th} residue respectively. The generated dynamic cross correlation matrix was constructed in Origin software.

1.5.3 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a covariance-matrix-based mathematical technique that is able to demonstrate atomic displacement and the loop dynamics of a protein⁵⁰. Prior to processing the MD trajectories for PCA, the trajectories of the free enzyme (Apo) and the NITD008-bound complex (Complex) were stripped of solvent and ions using the PTRAJ module in AMBER 14. The stripped trajectories were then aligned against their corresponding fully minimized structures. PCA was performed for C- α atoms on 900 snapshots each. Using in-house scripts, the first two principal components were calculated and the covariance matrices were generated. The first two principal components (PC1 and PC2) generated from each trajectory were averaged for both the free-enzyme and NITD008-complex. The first two principal components (PC1 and PC2) were computed and a 2 X 2 covariance matrix were generated using Cartesian coordinates of C α atoms. PC1 and PC2 correspond to first two eigenvectors of covariant matrices. Origin software⁵¹ was used to construct PCA plots.

Results and Discussion

3.1 NITD008-NS3 Helicase Complex

3.1.1 Binding of NITD008 with ZIKV Helicase

Research into ZIKV inhibitors has been minimal before 2016. However, NITD008, a *flavivirus* adenosine analogue was evidenced, both *in vitro* and *in vivo*, to inhibit ZIKV replication. The adenosine nucleoside analogue competes with natural ATP substrates, which are incorporated into the growing RNA chain. By this substitution, NITD008 is incorporated into the RNA chain, thus terminating the RNA elongation and inhibiting ZIKV maturation³⁷.

Molecular docking has become a major computational tool that is used to predict the orientation of a ligand at a binding site on the receptor. Results from docking often display multiple predicted orientations of the ligand within the active pocket⁵².

In this study, NITD008 docked at the ATP-binding site in 6 favorable conformations (Figures S2-S6), with the highest binding-affinity being -8.2 kcal/mol. Scoring functions often attempt to reproduce experimental binding affinities, but most software do not always yield the best prediction. Validation of the docked structure with experimentally known drugs was also not possible due to the lack of FDA inhibitors against ZIKV^{45,53,54}.

In an attempt to improve the binding affinity prediction of NITD008, all 6 predicted complexes were subjected to 130 ns molecular dynamic simulations, allowing for more realistic receptor flexibility in an implicit solvent. Each complex was then analyzed using the accurate, MM/GBSA, free binding energy calculation to determine the most favorable pose of NITD008 at the NS3 ATPase active site^{47,55-57}.

3.1.2 Free Energy calculations

The total binding free energy for each of the 6 poses of the NITD008- NS3 helicase complex were calculated using the MM/GBSA approach to better understand the various energy contributions within the binding pocket and assess which binding pose would show the most favorable intermolecular interactions at the helicase active site. Per residue decomposition analysis was also assessed and the residue-ligand interaction network of each pose were depicted as “ligplot” maps (Figures S2-S6). Of the six systems, the pose with the highest docking score, -8.2 kcal/mol, showed the most favorable free binding energy (-55.90 kcal/mol) supported the molecular docking score, indicating a favorable structural pose of NITD008 at the binding site.

The thermodynamic energy contribution of NITD008 to the total binding free energy of the complex surmounts to the stability of NITD008 in the ATP binding pocket and thus the stability of the complex during the simulation. Table 1 summarizes the free binding energy of the system taking into account the energies of the NS3 helicase and NITD008.

Table 6.1: Summary of free binding Energy contributions to the NITD008-NS3 Helicase system.

Energy Components (kcal/mol)					
	ΔE_{vdw}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV HELICASE	-3429.35 ± 30.09	-28758.51 ± 159.37	-32187.86 ± 155.05	-5121.93 ± 115.09	-37309.79 ± 71.27
NITD008	-4.69 ± 0.85	18.12 ± 5.27	13.43 ± 5.28	-221.12 ± 3.35	-207.68 ± 3.72
COMPLEX	-37.71 ± 4.12	-382.94 ± 28.72	-420.64 ± 28.59	364.75 ± 22.80	-55.90 ± 7.71

Figure 6.2 represents the residue interaction plot of NITD008 within the active site. The active site residues Gly199, Lys200 and Glu286 formed stable hydrogen bonds with highly electronegative oxygen atoms of NITD008. The residues pocketing NITD008 within the active site included Gly197, Ala198, Gly199, Lys200, Thr201, Arg202, Glu288, Gly415, Asn417 and Arg456.

It was also interesting to note that the most favorable NITD008-pose shared five active residues with the ATP-bound helicase reported by Tian *et al* (2016). The crystal structure of the ATP-bound helicase showed Lys200 to stabilize the triphosphate of the ATP¹⁴. The Lys200 of the NITD008-bound helicase showed a similar stabilizing hydrogen bond with the terminal hydroxyl group located on the ribose of NITD008.

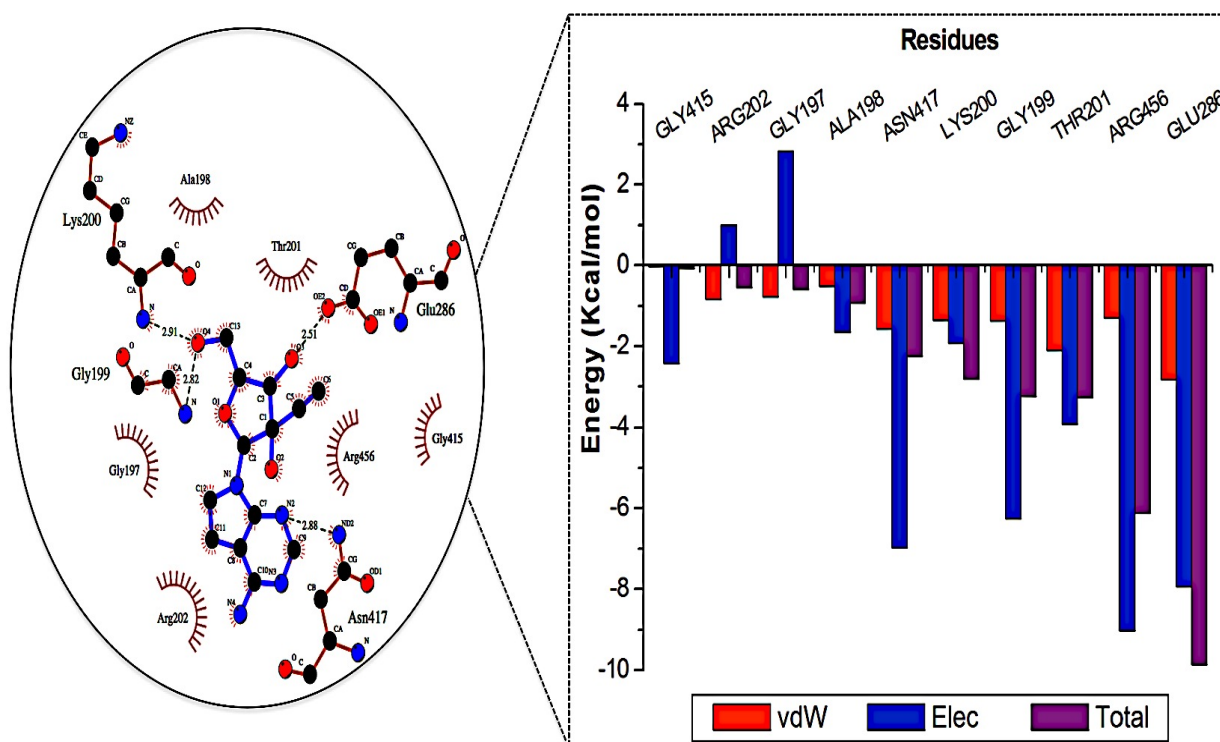


Figure 6.2: Energy contributions of the highest interacting residues at the ATPase active site. The residue ligand interaction network illustrates stabilizing hydrophobic interactions pocketing NITD008 at the active site. The highest energy contribution was a hydrogen bond interaction shared between Glu286 and the 3rd oxygen of the ribose component of NITD008.

Superimposition of NITD008-docked NS3 helicase with the ATP-NS3 helicase complex demonstrated both compounds to bind in a hydrophilic conformation despite the carbon and acetylene substitutions at N-7 of the purine and the 2' position of the ribose, respectively (Figure 6.3).

The structural similarities between NITD008 and ATP, as well as the active site residue interactions and accurate free-binding energy prompted the further analysis of NITD008-complex.

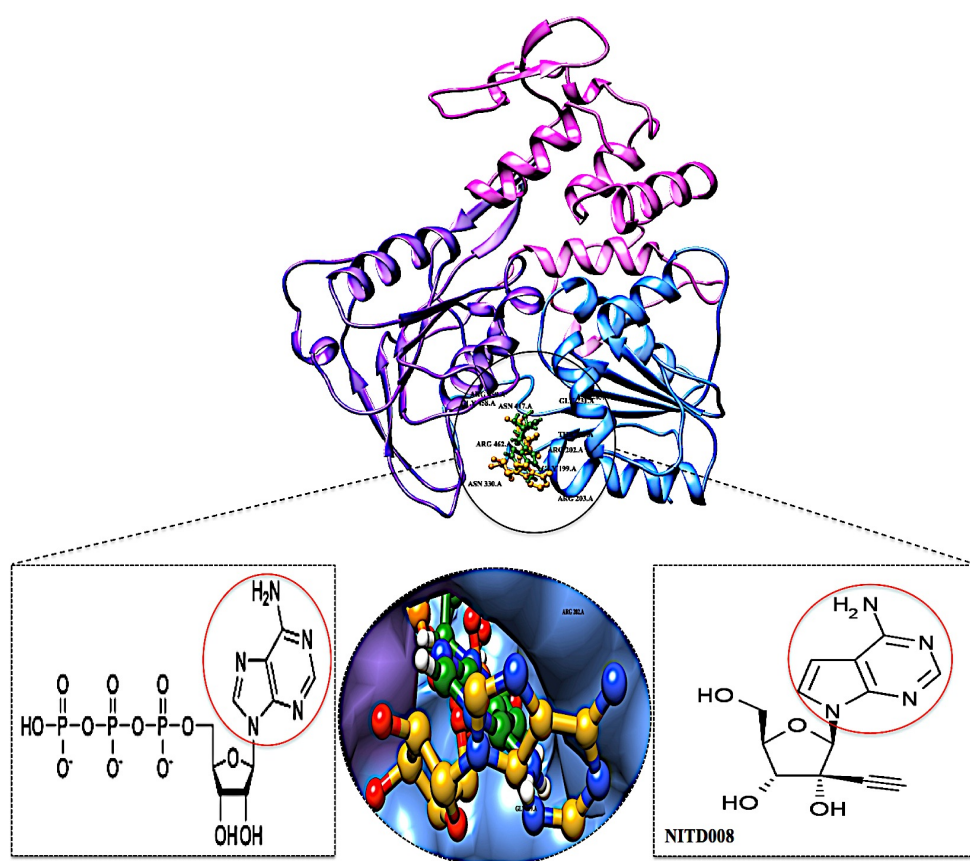


Figure 6.3: Superimposed conformation of structurally similar NITD008 and ATP docked at ATPase site of ZIKV NS3 Helicase.

3.2 Systems Stability

The length of a MD simulation is paramount when establishing insights into the structural dynamics of a biological system. With an extended simulation time, a system is able to reach convergence, thus becoming stable. To assure the equilibration of the simulation, the potential energy and temperature were monitored (Figure S1). The average potential energy (-145774 kcal/mol) was measured at 300K, suggesting a stable conformation at this temperature.

3.2.1 Stability of NS3 Helicase Apo and Bound System

The C- α backbone root mean square deviations (RMSD) were monitored throughout the 130 ns MD simulation for both the free (Apo) enzyme and the complex. Both systems reached convergence after 60 ns (RMSD deviation < 2 Å). It can be noted that the C- α backbone atoms in both systems stabilized after a 40 ns time period, although, fluctuations in rigidity did increase during the 47-52ns period in the NITD008 complex (Figure 6.4). This could possibly be due to the occurrence of conformational changes because of the bond interactions taking place between NITD008 and the active site residues as seen in the Per-residue energy decomposition.

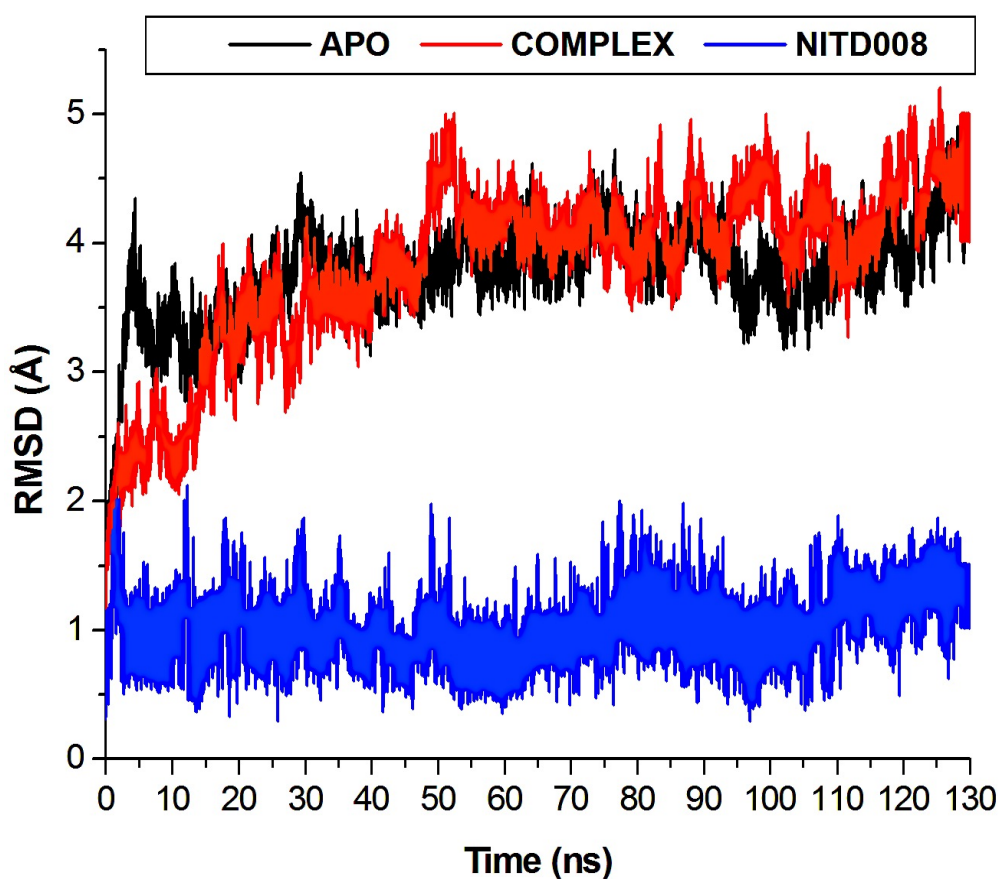


Figure 6.4: C- α backbone RMSD for NS3 Helicase Apo enzyme and NITD-complex conformation. The average C- α RMSD was calculated to be 3.62 Å and 3.77 Å, respectively. Increased fluctuations occurred at 47-52 ns in the NITD008-complex.

3.2.2 Conformational Fluctuations of the NS3 Helicase

To better understand the structural changes that may be occurring upon ligand binding, the root mean square fluctuation (RMSF) of the C- α atoms of each residue in the Apo system and NITD008-complex were calculated. Figure 6.5 clearly demonstrates greater flexibility of residues of the NITD008-complex when compared to the Apo enzyme. Fluctuations take place between residues 198-204, which form distinct hydrophobic and hydrogen bond interactions with NIT008D at the active site. This region, the P-Loop, is found in all *flavivirus* helicases and has been shown to have flexibility during binding of ATP¹⁴. The P-loop adopts structural modifications to accommodate the binding of ATP and Mn²⁺. This flexibility extends greatly in comparison to the Apo enzyme, thus verifying ZIKV P-loop flexibility upon ligand-binding. Other fluctuations occurred in domain II, and I around the ATP-active site, at residues 244-248 and 325-348.

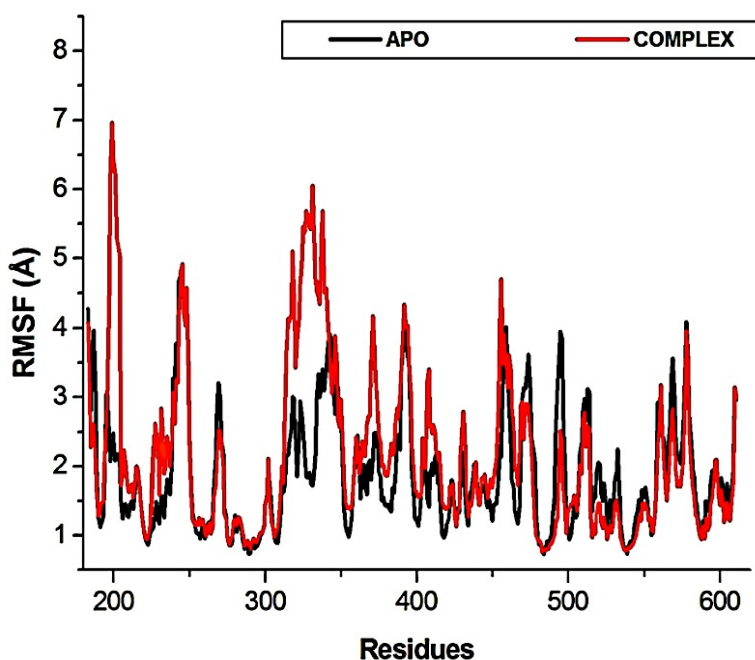


Figure 6.5: The RMSF of Apo enzyme and NITD008-complex. The structural flexibility in domain I and II is highly attributed to the binding of NITD008 to the ATP-active site. This is substantiated by the average RMSF of the NITD008-complex (2.17 Å), which is significantly higher than that of the Apo enzyme (1.90 Å).

3.2.3 Distribution of Atoms around the NS3 Helicase Backbone

The radius of gyration around the C- α atoms can measure the shape and folding of NS3 helicase before and after NITD008 binding. The radius of gyration measures the distribution of atoms from the center of mass (COM), thus indicating how compact a system is. Both the ApoA (22.05 Å) and NITD008 (22.17 Å) showed very similar structural compactness, however, there was an atomic distribution in the NITD008-complex from 40-58ns (Figure 6.6). This correlates with the escalated instability of the complex at 47-52ns demonstrated in the RMSD plot.

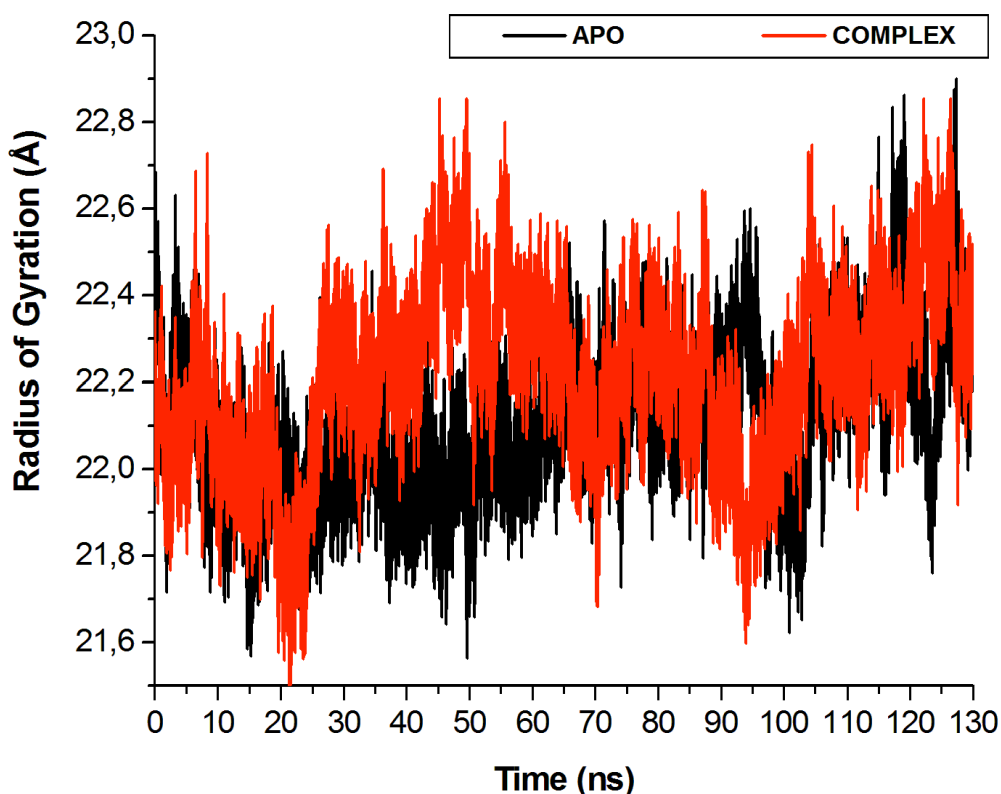


Figure 6.6: The radius of gyration (RoG) plot illustrating the difference in enzyme compactness of the NITD008-complex compared to the Apo enzyme.

The flexibility calculated from the RMSD, RMSF and RoG encouraged us to explore the dynamic structural modifications of the NS3 Helicase after NITD008 binding.

3.3 Investigation of the Dynamic Structural features ATP-Active Binding Region

3.3.1 Loop Flexibility and Distance metrics

The ZIKV NS3 Helicase is made up of three known flexible loops that are common to all *flaviviruses*: The P-loop (residues 196-203), the RNA-binding loop (residues 244-255) and the β -hairpin loop (residues 431-444). These loops may vary in size depending on the type of virus; however, they all have the same fundamental structural flexibility. The RMSF plot demonstrated major fluctuations at the P-loop as well as the RNA-binding loop, the β -hairpin loop however, showed no significant conformational change compared to the Apo enzyme. The plot also illustrated a flexible “325-338” region. Figure 6.7 depicts three snapshots of the Apo enzyme and NITD008-complex, taken at different intervals along the trajectory. Clear conformational shifts are illustrated along the trajectory in both Apo and bound systems.

To further investigate the conformational changes of the NS3 Helicase upon ligand binding, dynamic cross-correlation matrix (DCCM) analysis was performed at different conformational positions of the $C\alpha$ backbone atoms of the free protein and ligand-bound complex. Highly correlated motions of residues are represented in the red to yellow regions, whereas, the negative/anti-correlated movements of residue $C\alpha$ atoms are represented by blue-navy regions. It is evident from the correlation map that more globally correlated motion is observed in the case of the free protein, confirming conformational shifts after ligand binding. The latter residues of the NS3 Helicase, being residues 500-600, displayed anti-correlated movements in both the Apo and Bound complex, supporting the residue fluctuations in Figure 6.5. Figure 6.7 also depicts anti-correlation motions at residues “340-390”, which may be explained by the snapshots, in which, the flexible region in the NITD008-bound complex was converted from a $_{310}$ -helix to a α -helix.

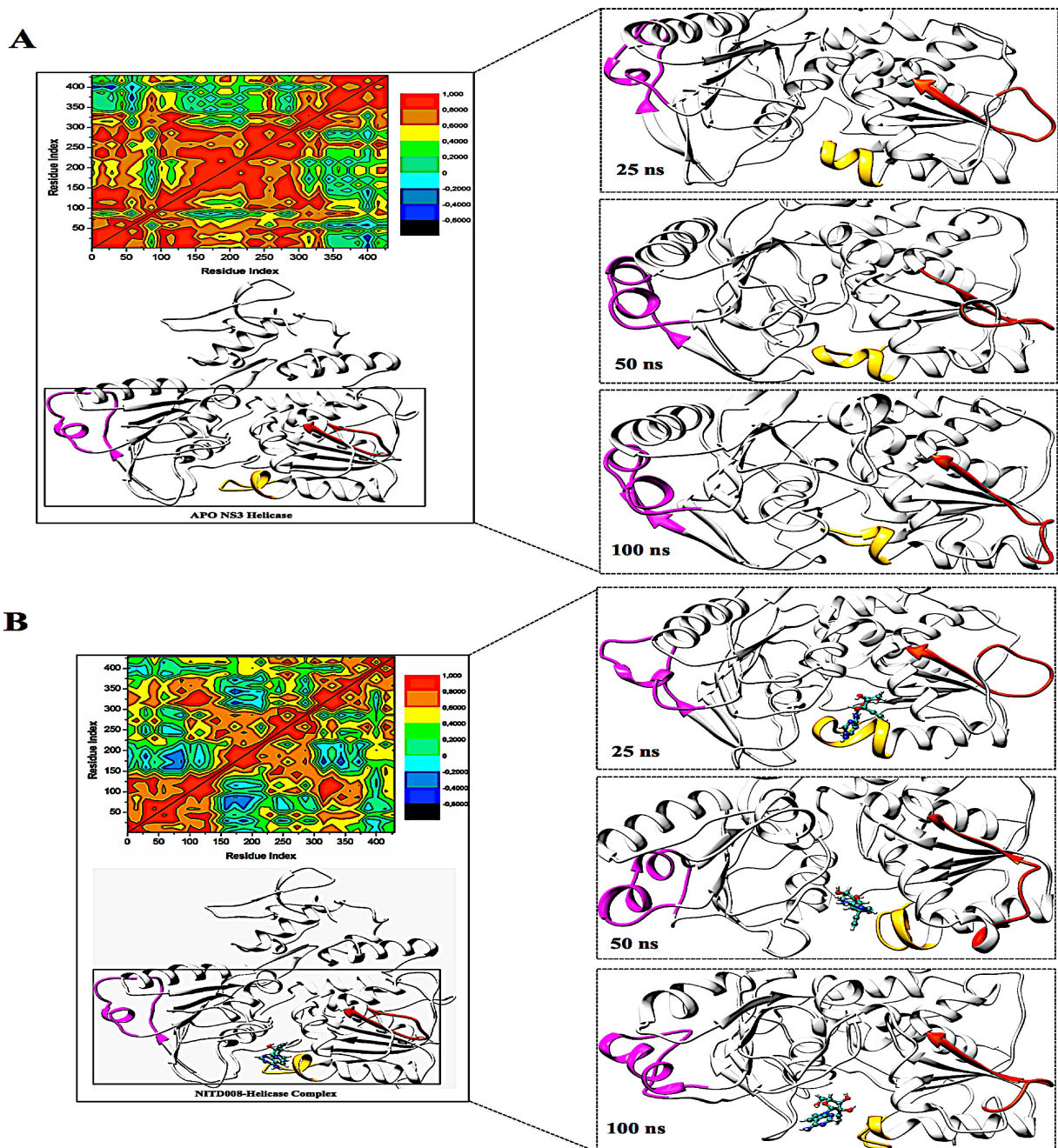
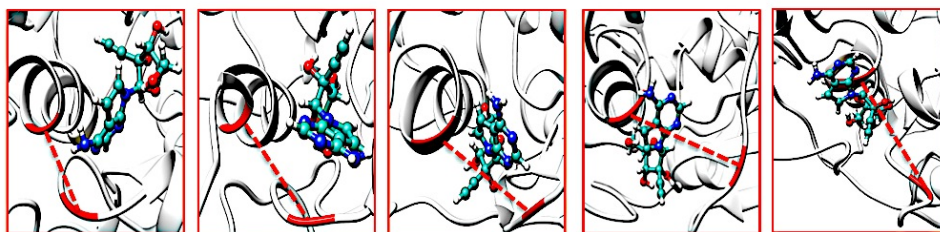


Figure 6.7: Structural Flexibility of the P-Loop (196-203), RNA-binding loop (244-255), and the $_{310}$ Helix (339-348) along the trajectory. The RNA-binding loop (orange) showed the loop shifting down in the Apo structure but an upward shift in the NITD008- Helicase complex. The P-Loop (Yellow) shifted away from the active site in the bound complex but closed in on the active site when no ligand was present. In the Apo structure, the helix-loop-helix stayed, with vibrational movement during the simulation, although, in the bound complex, the $_{310}$ Helix (Pink) was modified into a α -helix due to ligand motional shifts further into the hydrophobic pocket..

The P-loop clearly illustrates that when NS3 Helicase is in its Apo form and exposed to a 130 ns simulation, the P-loop closes on the active site by uncoiling the α -helix at Arg203 to form part of the loop. The loop tip (Ala198) and the adjacent catalytic residue (Gly451) had an average distance of 9.71 Å compared to the NITD008-complex distance of 12.75Å, whereby, as NITD008 becomes more stable at the active site and forms bond interactions, the P-loop is directed away from NITD008 and a larger catalytic space becomes available for the ligand as it forms stable hydrophobic interactions deeper within the hydrophobic pocket (Figure 6.8).

The distance between Ala198 and Gly451 of the Bound enzyme over the trajectory:



The distance between Ala198 and Gly451 of the APO enzyme over the trajectory:

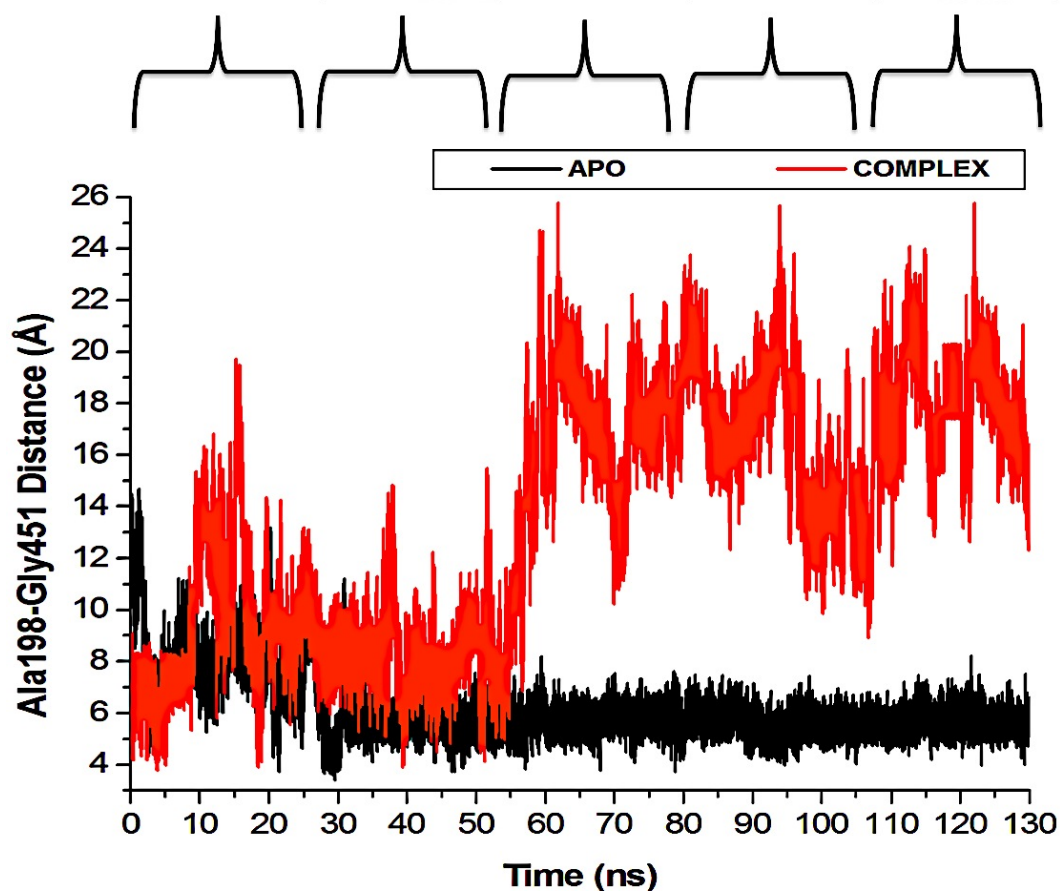
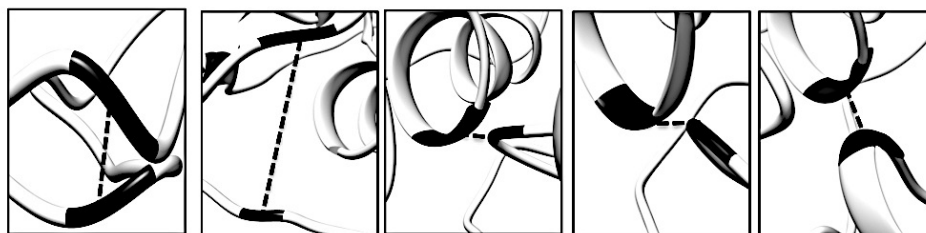
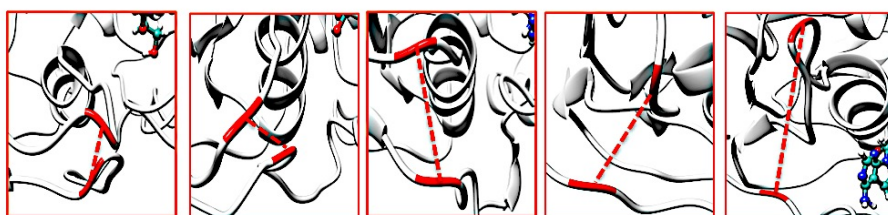


Figure 6.8: Residue fluctuations at the P-Loop region. The Apo enzyme illustrates closing of the loop at the active site due to a vacant hydrophobic pocket. Subsequent to ligand binding and the initiation of stabilizing hydrogen and hydrophobic bond interactions, the P-loop shifts down to accommodate the ligand, thus increasing the size of the hydrophobic pocket.

The “325-348” region demonstrates opposing conformational modifications between the Apo and complex systems compared to that of the P-loop. The Distance between the two catalytic residues from the loop tips; residue Ser324 and residue Asn448, measured for the Apo and NITD008-complex was 6.34 Å and 8.34 Å, respectively (Figure 6.9). The NITD008-complex had a greater distance between the residues due to the unraveling of 2 β -sheets found in domain II. This led to a “325-338” loop shift behind the active site and the “339-348” region being modified from a $_310$ Helix to a α -Helix (Figure 6.7). The $_310$ Helix conversion could be due to many reasons including changes in pH, interactions with other proteins and in this case, ligand binding. The ligand-protein interactions lead to distances between nitrogen and oxygen atoms from the protein backbone to fluctuate and as NITD008 moved further into the hydrophobic pocket, these fluctuations and hydrogen bond conversions caused the $_310$ helix to convert to an α -helix. These changes are important in illustrating the conformational fluctuations upon ligand binding.

The distance between Ser324 and Asn458 of the Bound enzyme over the trajectory:



The distance between Ser324 and Asn458 of the APO enzyme over the trajectory:

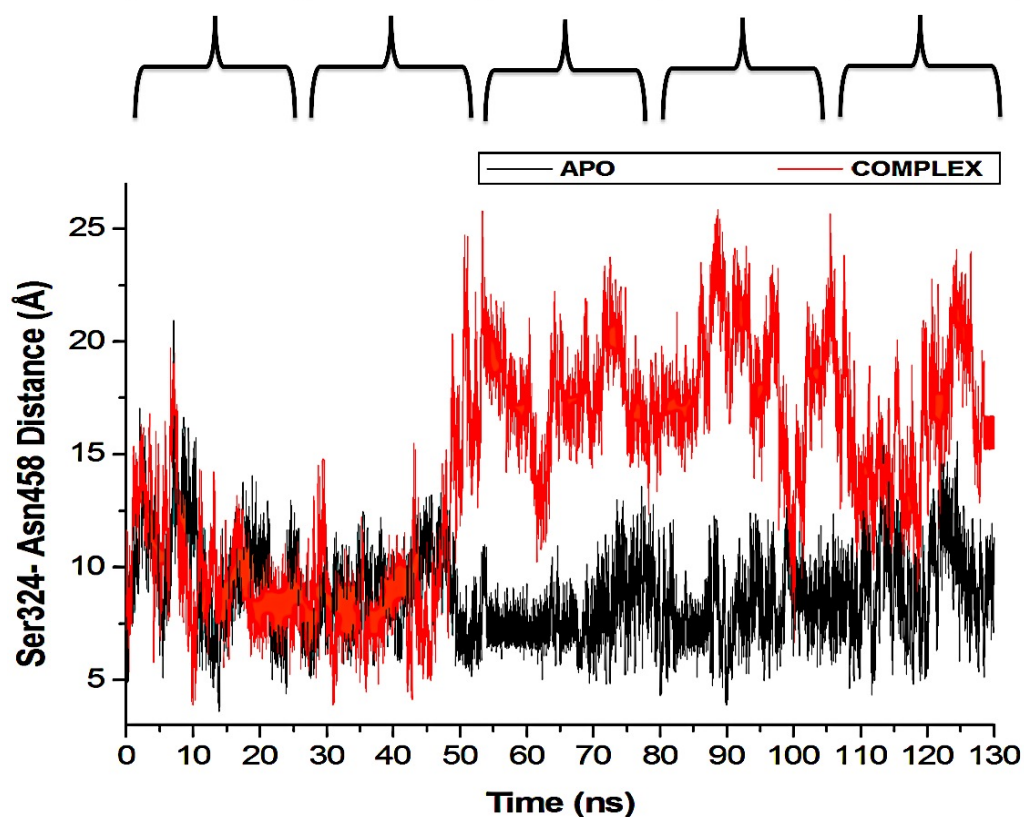
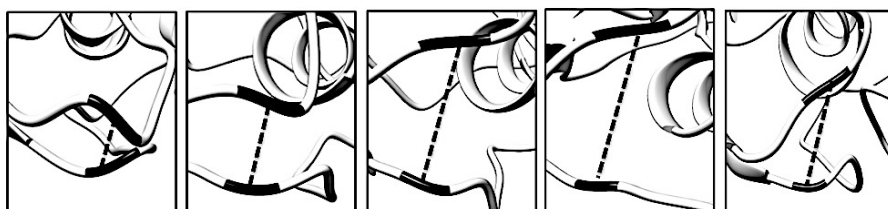


Figure 6.9: Residue fluctuations at the “325-348” region. The Apo enzyme illustrates widening of the loops of the Apo enzyme. The rear loop shifts down as the P-loop closes in on the active site. The largest fluctuation is seen after system stabilization at 40-60 ns. The NITD008-Helicase enzyme shows instability in both loops throughout the simulations, although, there was no widening of the loops as the rear loop shifted back rather than downward movement seen in the Apo system.

3.3.2 Principal Component Analysis

Conformational transitions of the free protein and NITD008-bound complex were characterized using PCA, a technique that has been widely employed to present experimentally detected conformational variations. Figure 6.10 highlights the motional shifts across two principle components in the case of NITD008-bound and unbound NS3 Helicase. It is evident that eigenvectors computed from the respective simulations varied immensely between the two systems, further elaborating on the dynamic conformational fluctuations from free to ligand-bound protein. The unbound system shows restricted structural motions of residue C α atoms, whereby the NITD008-bound system shows a larger spatial occupancy, thus substantiating the rigidity of the unbound system. This corresponds with the stability of the systems, illustrating greater distribution of the atoms around the center of mass and the system stability deviations for the NITD008-bound system. Correlation from analysis of both the free and bound protein demonstrates structural loop flexibility after binding of NITD008 to the ATPase active site.

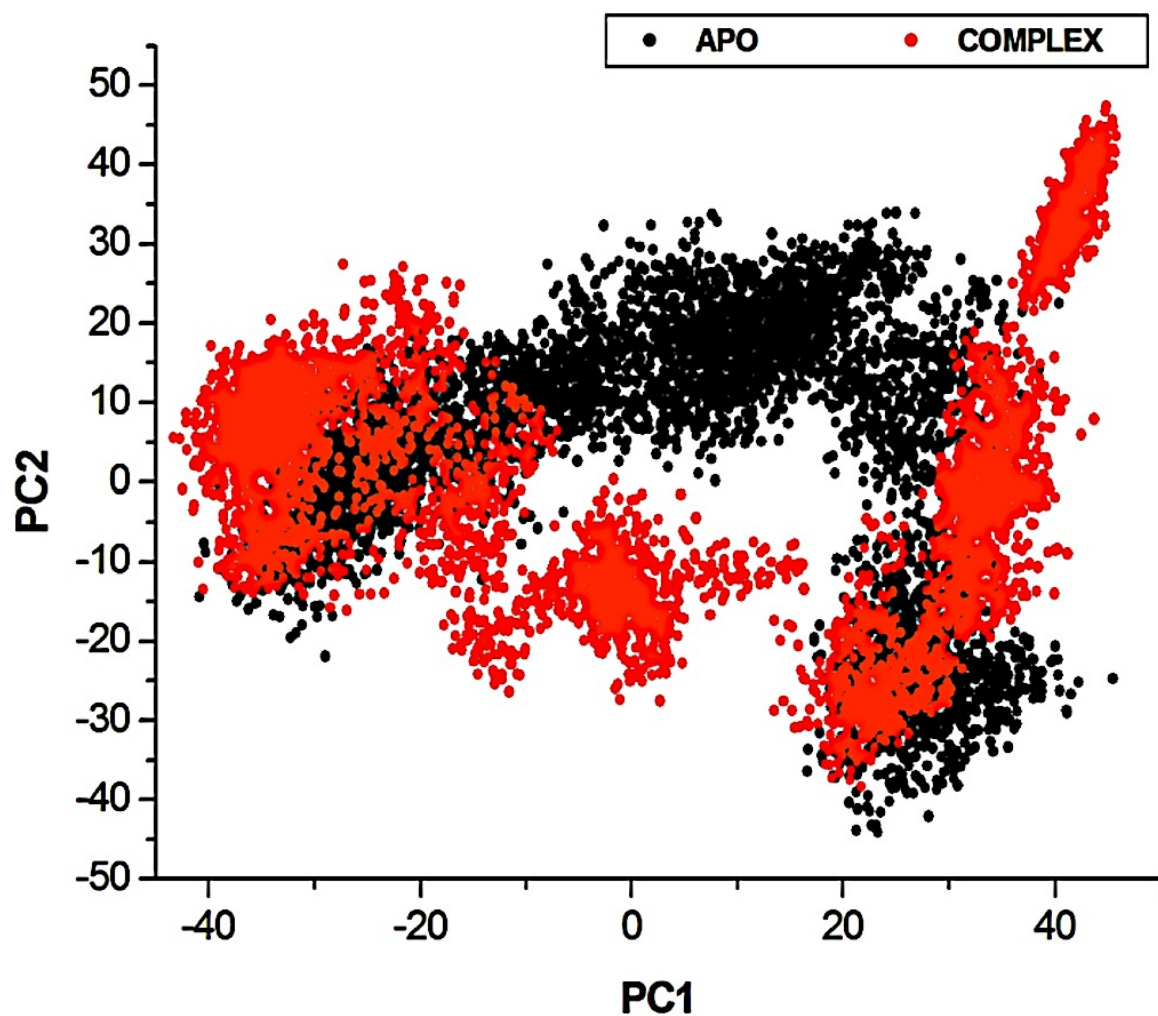


Figure 6.10: Projection of Eigen values of the Ca backbone, during 130 ns simulation, for Apo and NITD008-bound conformations of NS3 Helicase along the first two principal components. The X- and Y- axis, PC1 and PC2, respectively, represent a covariance matrix after elimination of eigenvectors (rotational movements). Each point between the single-directional motions represents a unique conformation during the simulation, whereby, similar structural conformations overlap in the graph.

Conclusion

The detailed MD analyses provided in this report demonstrate the structural alterations in ZIKV NS3 Helicase loop flexibility subsequent to binding of potent inhibitor, NITD008³⁷. Molecular simulations revealed profound motional shifts of the ZIKV P-Loop at the ATPase active site. This flexibility was revealed in the RMSF analysis and verified by graphical investigation of the loop at different time intervals during the simulation. Investigation into the dynamic cross-correlation of the unbound and bound systems as well as a plot of conformational poses along the first two principal components resulted in strongly significant structural flexibility of the NITD008-NS3 Helicase system compared to the rigid unbound protein. The P-loop has demonstrates similar motional shifts in other *flaviviruses* as well as in ZIKV, when natural substrate, ATP binds at the active site. The competitive inhibitor, NITD008, has been proven to effectively constrain ZIKV replication both *in vitro* and *in vivo*. Complex stability measured through the 130 ns simulation showed consistency of NITD008 at the ATPase active site and binding free energy calculations and residue-ligand networks revealed strong stabilizing hydrophobic and hydrogen bond interactions pocketing NITD008 in the active site. Further conformational changes were illustrated by the “325-338” loop shift behind the active site and the “339-348” region being modified from a fluctuating $_{310}$ Helix to a more stable α -Helix.

Crystallographic studies have identified the P-loop, specifically Lys200, to be critical in stabilizing the triphosphate moiety of an NTP, thus allowing flexibility upon ligand binding and activation¹²⁻¹⁴. To augment these key findings, Lys200 showed strong hydrogen bonds with the NTP-analogue, NITD008. Other active-hotspot residues included P-loop residues: Gly197-Arg202, Ala198, Glu286, Gly415, Asn417 and Arg456. The insights demonstrating the above binding landscape of the ZIKV NS3 Helicase will aid researchers in the identification of targeted-small molecule inhibitors through structure based drug design and to utilize pharmacophore models in screening for effective drugs with minimal toxicity.

Future experimental analysis is needed to fully understand these loop shifts toward inhibition of the enzyme as well as investigations into possible mutational resistance as seen in other *flavivirus* Helicase NTPase sites.

Acknowledgements

The authors acknowledge the National Research Foundation for their financial support (UID: 102103) and the Center for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

Notes and References

- 1 L. Broxmeyer and R. Kanjhan, *Mod. Res. Inflamm.*, 2016, **5**, 20–30.
- 2 Centers for Disease Control, *Centers Dis. Control Prev. Zika Virus Home*, 2016, 1–12.
- 3 S. A. Rasmussen, D. J. Jamieson, M. A. Honein and L. R. Petersen, *N. Engl. J. Med.*, 2016, **374**, 1981–1987.
- 4 A. M. Palomo, *J. Public Health Policy*, 2016, **37**, 133–135.
- 5 WHO, 2016, 1–12.
- 6 O. Faye, C. C. M. Freire, A. Iamarino, O. Faye, J. V. C. de Oliveira, M. Diallo, P. M. Zanutto and A. A. Sall, *PLoS Negl. Trop. Dis.*, 2014, **8**, 1–10.
- 7 R. Tilak, S. Ray, V. W. Tilak and S. Mukherji, *Med. J. Armed Forces India*, 2016, **72**, 157–163.

- 8 C. G. Noble, Y. L. Chen, H. Dong, F. Gu, S. P. Lim, W. Schul, Q. Y. Wang and P. Y. Shi, *Antiviral Res.*, 2010, **85**, 450–462.
- 9 M. Mahfuz, A. Khan, H. Al Mahmud, M. Hasan, A. Parvin, N. Rahman and S. M. B. Rahman, *Indian J. Pharm. Biol. Res.*, 2014, **2**, 44–57.
- 10 C. Zanluca, C. N. Duarte and D. Santos, *Microbes Infect.*, 2016, **18**, 295–301.
- 11 P. Ramharack and M. E. S. Soliman, *RSC Adv.*, 2016, **6**, 68719–68731.
- 12 R. Jain, J. Coloma, A. Garcia-Sastre and A. K. Aggarwal, *Nat. Struct. Mol. Biol.*, 2016, **2**, 1–4.
- 13 H. Tian, X. Ji, X. Yang, W. Xie, K. Yang, C. Chen, C. Wu, H. Chi, Z. Mu, Z. Wang and H. Yang, *Protein Cell*, 2016, **7**, 450–454.
- 14 H. Tian, X. Ji, X. Yang, Z. Zhang, Z. Lu, K. Yang, C. Chen, Q. Zhao, H. Chi, Z. Mu, W. Xie, Z. Wang, H. Lou, H. Yang and Z. Rao, *Protein Cell*, 2016, **7**, 562–570.
- 15 X. Cao, Y. Li, X. Jin, Y. Li, F. Guo and T. Jin, *Nucleic Acids Res.*, 2016, **44**, 10505–10514.
- 16 E. D’Ortenzio, S. Matheron, X. de Lamballerie, B. Hubert, G. Piorkowski, M. Maquart, D. Descamps, F. Damond, Y. Yazdanpanah and I. Leparac-Goffart, *N. Engl. J. Med.*, 2016, **374**, 2195–2198.
- 17 A. C. Gourinat, O. O. Connor, E. Calvez, C. Goarant and D.-R. M., *Emerg. Infect. Dis.*, 2015, **21**, 84–86.
- 18 M. J. Turmel, M. J. P. Hubert, Y. M. V Maquart, M. Guillou-Guillemette and I. Leparac-Goff, *Lancet*, 2016, **6736**, 2501–2501.
- 19 A. R. Plourde and E. M. Bloch, *Emerg. Infect. Dis.*, 2016, **22**, 1–15.

- 20 J.-M. Anaya, C. Ramirez-Santana, I. Salgado-Castaneda, C. Chang, A. Ansari, M. E. Gershwin, R. Martines, J. Bhatnagar, M. Keating, L. Silva-Flannery, A. Muehlenbachs, J. Gary, C. Woods, A. Parker, B. Wakerley, A. Uncini, N. Yuki, J. Anaya, Y. Shoenfeld, A. Rojas-Villarraga, R. Levy, M. Dalakas, B. Wakerley, N. Yuki, S. Kivity, M. Arango, M. Ehrenfeld, O. Tehori, Y. Shoenfeld, J. Anaya, A. Denman, B. Rager-Zisman, T. Kolter, H. Willison, N. Yuki, R. Lardone, N. Yuki, F. Irazoqui, G. Nores, I. Kostovic, R. Ghiulai, M. Sarbu, Z. Vukelic, C. Ilie, A. Zamfir, T. Bell, E. Field, H. Narang, O. Faye, C. Freire, A. Iamarino, O. Faye, J. Oliveira, M. Diallo, A. Nahmias, S. Nahmias and D. Danielsson, *BMC Med.*, 2016, **14**, 1–3.
- 21 N. L. Bayless, R. S. Greenberg, T. Swigut, J. Wysocka and C. A. Blish, *Cell Host Microbe*, 2016, **20**, 423–428.
- 22 D. Olagnier, M. Muscolini, C. B. Coyne, M. S. Diamond and J. Hiscott, *DNA Cell Biol.*, 2016, **35**, 367–372.
- 23 J. B. Brault, C. Khou, J. Basset, L. Coquand, V. Fraissier, M. P. Frenkiel, B. Goud, J. C. Manuguerra, N. Pardigon and A. D. Baffet, *EBioMedicine*, 2016, **10**, 71–76.
- 24 H. Li, L. Saucedo-Cuevas, J. A. Regla-Nava, G. Chai, N. Sheets, W. Tang, A. V. Terskikh, S. Shresta and J. G. Gleeson, *Cell Stem Cell*, 2016, **19**, 593–598.
- 25 T. J. Nowakowski, A. A. Pollen, E. Di Lullo, C. Sandoval-Espinosa, M. Bershteyn and A. R. Kriegstein, *Cell Stem Cell*, 2016, **18**, 591–596.
- 26 J. Cohen, *Science (80-.)*, 2016, **351**, 543–544.
- 27 E. Kim, G. Erdos, S. Huang, T. Kenniston, L. D. Falo and A. Gambotto, *EBioMedicine*, 2016, **13**, 315–320.
- 28 T. C. Pierson and B. S. Graham, *Cell*, 2016, 167, 625–631.

- 29 G. W. A. Dick, S. F. Klitchen and A. J. Haddow, *Trans. R. Soc. Trop. Med. Hyg.*, 1969, **63**, 708–737.
- 30 R. W. Malone, J. Homan, M. V Callahan, J. Glasspool-Malone, L. Damodaran, A. D. B. Schneider, R. Zimler, J. Talton, R. R. Cobb, I. Ruzic, J. Smith-Gagen, D. Janies, J. Wilson, D. Hone, S. Hone, S. Bavari, V. Soloveva and S. Weaver, *PLoS Negl. Trop. Dis.*, 2016, **10**, 1–26.
- 31 A. N. Hazin, A. Poretti, D. Di Cavalcanti Souza Cruz, M. Tenorio, A. van der Linden, L. J. Pena, C. Brito, L. H. V. Gil, D. de Barros Miranda-Filho, E. T. de A. Marques, C. M. Turchi Martelli, J. G. B. Alves and T. A. Huisman, *N. Engl. J. Med.*, 2016, **374**, 2193–2195.
- 32 N. Gruba, J. I. Rodriguez Martinez, R. Grzywa, M. Wysocka, M. Skorenski, M. Burmistrz, M. Lecka, A. Lesner, M. Sienczyk and K. Pyrc, *FEBS Lett.*, 2016, **590**, 3459–3468.
- 33 B. D. Cox, R. A. Stanton and R. F. Schinazi, *Antivir. Chem. Chemother.*, 2016, **24**, 118–126.
- 34 J. Lei, G. Hansen, C. Nitsche, C. D. Klein, L. Zhang and R. Hilgenfeld, *Science (80-.)*, 2016, **353**, 503–5.
- 35 D. Luo, T. Xu, R. P. Watson, D. Scherer-Becker, A. Sampath, W. Jahnke, S. S. Yeong, C. H. Wang, S. P. Lim, A. Strongin, S. G. Vasudevan and J. Lescar, *EMBO J.*, 2008, **27**, 32090–3219.
- 36 A. V. Chernov, S. A. Shiryaev, A. E. Aleshin, B. I. Ratnikov, J. W. Smith, R. C. Liddington and A. Y. Strongin, *J. Biol. Chem.*, 2008, **283**, 17270–17278.
- 37 Y. Q. Deng, N. N. Zhang, C. F. Li, M. Tian, J. N. Hao, X. P. Xie, P. Y. Shi and C. F. Qin,

Open Forum Infect. Dis., 2016, **3**, 1–4.

- 38 H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, E. Philip, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, N. Thanki, H. Weissig and J. D. Westbrook, *Biol. Crystallogr.*, 2002, **58**, 899–907.
- 39 S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang and S. H. Bryant, *Nucleic Acids Res.*, 2016, **44**, 1202–1213.
- 40 S. Kusumaningrum, E. Budianto, S. Kosela, W. Sumaryono and F. Juniarti, *J. Appl. Pharm. Sci.*, 2014, **4**, 47–53.
- 41 Z. Yang, K. Lasker, D. Schneidman-Duhovny, B. Webb, C. C. Huang, E. F. Pettersen, T. D. Goddard, E. C. Meng, A. Sali and T. E. Ferrin, *J. Struct. Biol.*, 2012, **179**, 269–278.
- 42 S. Cosconati, S. Forli, A. L. Perryman, R. Harris, D. S. Goodsell and A. J. Olson, *Expert Opin. Drug Discov.*, 2010, **5**, 597–607.
- 43 M. F. Sanner, *Scripps Res. Inst.*, 2008, **26**, 1–12.
- 44 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 445–461.
- 45 N. Huang, B. K. Shoichet and J. J. Irwin, *J. Med. Chem.*, 2012, **49**, 6789–6801.
- 46 P. C. Nair and J. O. Miners, *Silico Pharmacol.*, 2014, **2**, 1–4.
- 47 M. Ylilauri and O. T. Pentikäinen, *J. Chem. Inf. Model.*, 2013, **53**, 2626–2633.
- 48 T. Hou, J. Wang, Y. Li and W. Wang, *J. Chem. Inf. Model.*, 2011, **51**, 69–82.
- 49 V. Gosu and S. Choi, *Sci. Rep.*, 2014, **4**, 1–13.
- 50 A. M. Martinez and A. C. Kak, *Trans. Pattern Anal. Mach. Intell.*, 2001, **23**, 228–233.

- 51 E. Seifert, *J. Chem. Inf. Model.*, 2014, **54**, 1552.
- 52 D. Ramirez and J. Caballero, *Int. J. Mol. Sci.*, 2016, **17**, 1–15.
- 53 X.-Y. Meng, H.-X. Zhang, M. Mezei and M. Cui, *Curr. Comput. Aided. Drug Des.*, 2011, **7**, 146–57.
- 54 L. G. Ferreira, R. N. Dos Santos, G. Oliva and A. D. Andricopulo, *Molecular docking and structure-based drug design strategies*, 2015, vol. 20.
- 55 P. A. Greenidge, C. Kramer, J. C. Mozziconacci and R. M. Wolf, *J. Chem. Inf. Model.*, 2013, **53**, 201–209.
- 56 J. M. Hayes and G. Archontis, *InTech*, 2011, 171–190.
- 57 F. Godschalk, S. Genheden, P. Söderhjelm and U. Ryde, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7731–9.

CHAPTER 7

Characterizing the Ligand Binding Landscape of Zika NS3 Helicase- Promising Lead Compounds as Potential Inhibitors

Sofiat Oguntade ^a, Pritika Ramharack ^a and Mahmoud E. S. Soliman ^{a*}

^aDiscipline of Pharmaceutical Sciences, School of Health Sciences, University of KwaZulu-Natal,
Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

1. Dean and Head of School of Health Sciences, Full Professor: Pharmaceutical Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.
2. Department of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt.
3. College of Pharmacy and Pharmaceutical Sciences, Florida Agricultural and Mechanical University, FAMU, Tallahassee, Florida 32307, USA.

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

Abstract

Aim: This study aims to provide insight into the binding features of the ATPase and ssRNA sites of the NS3 helicase. **Methods:** Clinically approved *flavivirus* inhibitors were docked to the corresponding active sites of the protein and the three best compounds were validated with molecular dynamic simulations. **Result:** Binding of Ivermectin to ssRNA site and Lapachol and HMC-HO1 α to the ATPase site allowed for conformational rigidity of the Zika NS3 helicase, thus stabilizing residue fluctuations and allowing for protein stability. Favorable free binding energies were also noted between compounds and the helicase, thus supporting the intermolecular forces at the helicase active site. **Conclusion:** The pharmacophoric characteristics found in Lapachol, HMC-HO1 α and Ivermectin may be utilized in the design of a potent hybrid drug that is able to show efficient inhibition of a multitude of diseases including the detrimental co-infection of ZIKV, Dengue and Chikungunya.

Keywords:

Zika virus, NS3 helicase, molecular dynamic simulations, free binding energy

1. Introduction

Zika virus (ZIKV) is a positive-sense, single stranded RNA arbovirus belonging to the genus *flavivirus* and family *flaviviridae* [1]. The virus was first discovered in a forest in Uganda called the Zika forest near lake victoria in 1947, thus coining the virus's name [2,3]. The virus was then isolated in the blood of a sentinel Rhesus monkey during research on the Yellow fever virus [4], while a second isolation was done in 1948 at the same site [5]. ZIKV virus has a wide geographical distribution including Africa (Uganda, Egypt, Gabon), Asia (India, Malaysia, Vietnam, Thailand, Indonesia), and Micronesia [6]. This has been demonstrated through viral isolations and serologic studies [7,8]. Although isolations of the virus were analyzed, researchers only detected the virus in humans in 1952 when neutralizing antibodies were picked up in infected sera. Scientists Boorman and Porterfield subsequently studied the transmission of viruses from mosquito to primates and based on further isolations from both mosquito and monkey concluded that mosquitoes acted as vectors for ZIKV [1].

The rapid spread of the virus across continents is primarily due to vector transmission via the *Aedes aegypti*, *Aedes albopictus* and *Aedes africanus* mosquito [2]. These Vectors are endemic to tropical and sub-tropical areas. However, due to evolving climates, the mosquitoes have expanded their habitat, thus increasing the number of mosquitoes as vectors of *flaviviruses* [9–11]. However, other routes of transmission have been reported, including, sexual transmission [10,12], perinatal transmission, and blood transfusion [13]. The symptoms following ZIKV viral infection are mild headache, maculopapular rash, fever, malaise, conjunctivitis and arthralgia. These symptoms are shared with other related *flaviviruses*, including Dengue virus, Yellow fever virus, West Nile, St. Louis encephalitis virus and Japanese Encephalitis virus [15,16]. The most recent and devastating outbreak of ZIKV occurred in Brazil, at the end of 2015. The virus has, to date, rampaged South America by being evidenced as a leading cause of microcephaly by prenatal transmission [17]. Increasing scientific evidence now shows that the virus is able to pass

through the blood-brain-barrier and infect neural cells, thus playing a role in diseases such as microcephaly and Gullian-Barré Syndrome [18].

The ZIKV genome contains 10.7kb single stranded RNA, which contains a large polyprotein, which cleaves into 3 structural proteins (envelope, E; membrane precursor, PrM; and capsid, C) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5), of which, the NS3 helicase plays a pivotal role in viral replication and RNA synthesis. Presently, researchers are focusing on the structural and non-structural viral proteins for the development of drugs [19], due to their crucial characteristics in viral replication [20]. The NS3 helicase (Figure 7.1) has three domains and two binding sites, being the adenosine triphosphate (ATP) and single stranded ribonucleic acid (ssRNA) site [21,22]. Inhibiting both the ATP and ssRNA sites will be crucial in the inhibition of the NS3 helicase as studies have shown that each domain may act independently from the other [23]. However, the close proximity of the two binding sites can bring about the possibility of designing a single inhibitor that can span both sites [3].

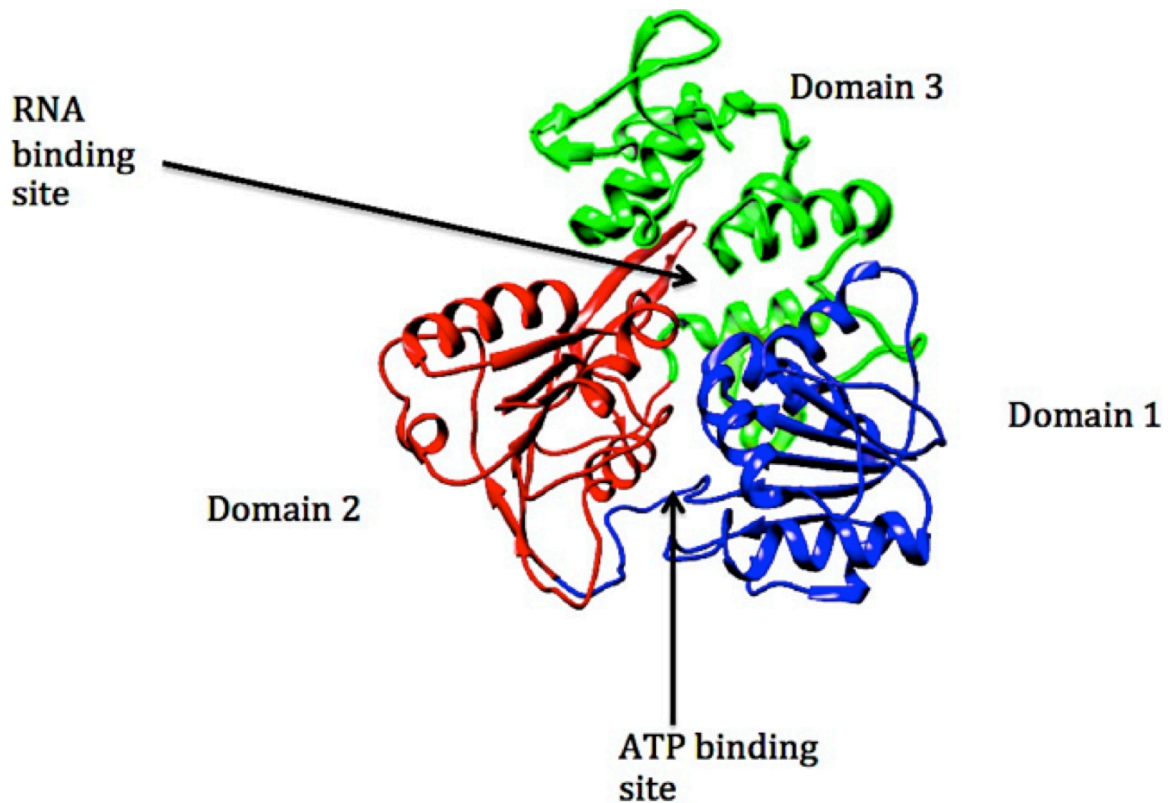


Figure 7.1: Structure of NS3 ZIKV Helicase (PBD 5JMT), [21]. Domain 1 (blue: residue 175-332) and domain 2 (red: residue 333-481) are seen facing each other and domain 3 (green: residue 482-617) lying above the other 2 domains. The ATP binding site is located in the cleft between domain 1 and domain 2 and the ssRNA binding site is located at the tunnel that separates domain 3 from the other 2 domains [22].

Due to the rapid spread of the disease on a global scale and the detrimental long-term complications, researchers such as Barrows *et al* (2016) and Xu *et al* (2016) have turned to ‘repurposing’ *flavivirus* FDA approved drugs rather than the lengthy process of designing and synthesizing new drugs [4,5]. One of the most widely used antihelminthic drugs, Ivermectin, has been evidenced to have potent inhibitory effects on *flaviviruses* by acting as a competitive inhibitor of viral ssRNA at the RNA binding site of the NS3 helicase [6,7]. Barrows *et al* (2016) validated Ivermectin as a potent ZIKV inhibitor in an *in vitro* screening study, alongside 17 other

FDA-approved *flavivirus* drugs as well as daptomycin, which had no previous anti-viral activity [4].

Another potential drug candidate against ZIKV is the adenosine nucleoside analog, NITD008, which has been reported to have competitive inhibitory properties against adenosine substrates *in vitro* and *in vivo* [8]. However, other reports have also shown elevated toxicity levels in preclinical animal testing [9].

One of the major challenges of ZIKV is its ability to co-infect the host. Multiple cases reporting Chikungunya, Dengue and Zika co-infection have been identified, leading to potentially exacerbated neurological effects on the host and fetus [10]. By this end, identifying potential inhibitors against ZIKV that have already been approved as a Dengue or Chikungunya treatment would be beneficial as it would be less toxic than administering multiple drugs to a patient [11].

Although numerous studies have been released elucidating ZIKV drug discovery, no FDA approved drugs are presently available. There is also a lack in literature regarding the structural and conformational features of the protein, thus designing effective novel small drug molecule inhibitors may be challenging.

In this study, we have utilized clinically approved *flavivirus* NS3 small molecule inhibitors to analyze the binding affinity and stability of the ZIKV NS3 domains via Molecular Dynamics (MD) simulations, thus mapping out binding hotspots and landscaping interactions of the complexes.

In addition to this, we will employ Accelerated Molecular Dynamics (aMD) in order to run the simulation for a longer time frame to ensure sufficient conformational sampling and accurate physical force field. Accelerated MD is an enhanced sampling technique that operates by

modifying potential energy, reducing the height of local barriers and accelerating transition between different low energy states [25][26]. This will enable the sampling of distinct bimolecular conformations and rare barrier-crossing events that cannot be easily accessed in a conventional MD simulation, thereby improving the efficiency of convectional MD [27].

2. Computational methodology

2.1 Protein structure preparation

The crystal structure of the *Escherichia coli* strain of Zika virus NS3 helicase was retrieved from protein data bank (PDB: 5JMT). It was then prepared for molecular docking by stripping it off of water molecules using UCSF CHIMERA [28] and adding the necessary hydrogen atoms using Molegro Molecular Viewer (MMV) [29].

2.2 Molecular docking

Molecular docking was performed on 10 ligands: 6 Naphthoquinones (Lapachol, Atovaquone, Parvaquone, Buparvaquone, α -Lapachone, β -Lapachone)[30], 3 purine nucleoside analogues (1-(2'-deoxy- α -D-ribofuranosyl)imidazo[4,5-*d*]pyridazine-4,7(5*H*,6*H*)dione)(HMC-HO1 α), 1-(2'-*O*-methyl- α -D-ribofuranosyl)imidazo[4,5-*d*]pyridazine- 4,7(5*H*,6*H*)-dione(HMC-HO4) and 1-(β -D-ribofuranosyl)imidazo [4,5-*d*]pyridazine-4,7(5*H*,6*H*)-dione(HMC-HO5)) [31,32] and Ivermectin [33]. Each of the compounds were then downloaded from PubChem [34], converted to mol2 format and assessed using MMV to ensure that they display the correct bond angle and hybridization state. The 2D structures of the ligands are given in the supplementary material (Figure S1).

Docking was carried out with the Autodock Vina software [35]. Ivermectin was docked at the ssRNA binding site, while the rest were docked at the ATPase binding site. The grid box parameters for the 2 sites are given in Figure S2. Of the 10 ligands docked into the active site of NS3 protein structure, the best 3 complexes were chosen and subsequently subjected to accelerated-MD.

2.3 Molecular dynamic simulations

Molecular Dynamic (MD) simulations were performed on the 3 complexes using the graphics processor unit (GPU) version of the PMEMD engine provided with the AMBER 14 package [36,37]. The Antechamber module was used to generate atomic partial charges for the ligands using GAFF force field [38]. The protein was described using the FF14SB of the Amber force field [39]. The LEAP module in AMBER 14 was used to generate topologies for the system by adding protons and counter ions to neutralize the system [36]. Subsequently, the complexes were then solvated in a TIP3P [40] octahedron water box with a distance of 8 Å away from the water box edge. The Periodic boundary conditions were employed and the particle-mesh Ewald method (PME) in AMBER 14 was used to treat the long-range electrostatic interactions with a non-bonding cut-off distance of 10 Å.

Minimization of the systems were performed with a restraint potential of 500 kcal/mol Å² to treat the solute for 1000 steepest descent steps using the SANDER module of the AMBER 14 program, followed by 1000 steps of conjugate gradient minimization. The systems were then minimized over 1000 steps with unrestrained conjugate gradient. Gradually, the systems were heated from 0 to 300 K for 50 ps, such that the system maintained a fixed number of atoms and a fixed volume, that is, a canonical (NVT) ensemble.

The entire system was then equilibrated at 300 K with a 2 fs time step in the NPT ensemble for 500 ps, and Berendsen temperature coupling [39] was used to maintain a constant pressure at 1 bar. The SHAKE algorithm [40] was employed on all atoms so as to constrain the bonds of all hydrogen atoms. With no restraints imposed, an initial production run was performed for 10 ns in an isothermal isobaric (NPT) ensemble using a Berendsen barostat with a target pressure of 1 bar and a pressure-coupling constant of 2 ps. The systems were subsequently subjected to 10 ns of accelerated MD using a set of parameters calculated from the potential energy of the converged

system (Table 1). Coordinates were saved every 1 ps and the trajectories were analyzed every 1 ps using the PTRAJ module of AMBER 14. Each system was consequently subjected to post molecular dynamic analysis including root mean square fluctuation (RMSF), root mean square deviation (RMSD) and radius of gyration (RoG). Included in analysis was the ligand-residue profile [25]. Visualization of trajectories was conducted in Chimera [28], while the results were analyzed and plots were generated with aid of Origin software [41].

Table 7.1: Calculated parameters for running accelerated molecular dynamics.

	ethreshP (kcal/mol)	ethreshD (kcal/mol)	alphaP (kcal/mol)	alphaD (kcal/mol)
Lapachol System	-144728	9424.9	1404.4	355.2
Ivermectin System	-144600	9465.1	1425.2	355.2
HMC- HO1 α System	-144698	9455.3	1404.4	355.2

2.4 Thermodynamic calculations

Over the years, molecular mechanics/generalized-born surface area (MM/GBSA) method of binding free energy calculations have proved to be a practicable means of understanding the ligand-residue landscape binding in various biological macromolecules [42–45]. Therefore, MM/GBSA approach was employed to calculate the binding free energies of Ivermectin, Lapachol and HMC-HO1 α bound to NS3 helicase protein. To achieve this, 1000 snapshots were

extracted from each of the 20 ns trajectories. The following equation describes the calculations of binding free energy.

The term E_{gas} denotes the gas-phase energy that consists of the internal energy E_{int} , Coulomb energy E_{ele} , and the van der Waals energies E_{vdw} . E_{gas} was directly estimated from the FF14SB force field terms. The solvation free energy, G_{sol} , is estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G_{SA} . The non-polar solvation energy, G_{SA} , is determined from the SASA using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB} , contribution is estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature, respectively.

2.5. Per-residue energy decomposition analysis

Per-residue free energy decomposition was carried out in order to obtain the contribution of each residue to the total binding free energy profile between the inhibitors Ivermectin, Lapachol and HMC-HO1 α with the NS3 helicase protein. This was achieved using the MM/GBSA method in AMBER 14[25].

3. Results and Discussion

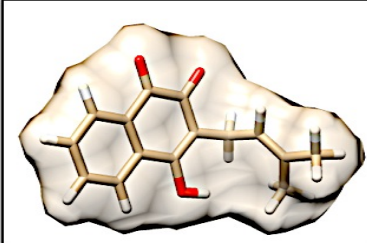
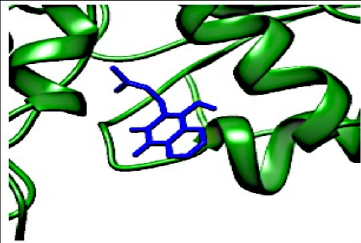
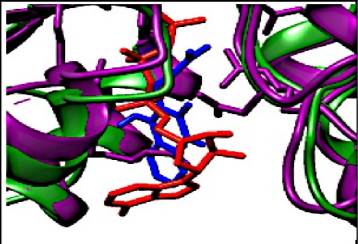
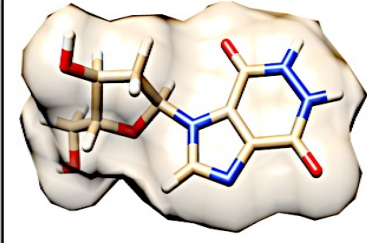
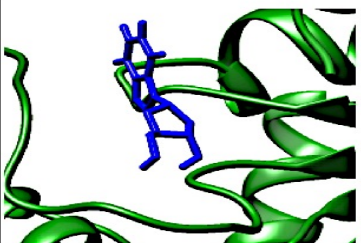

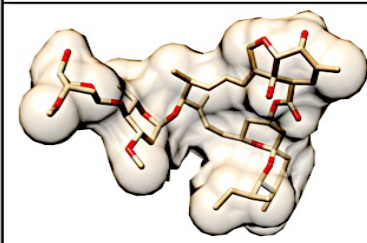
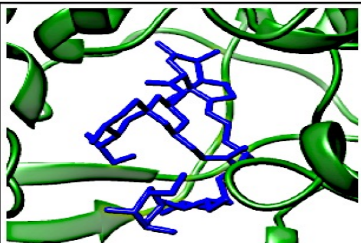
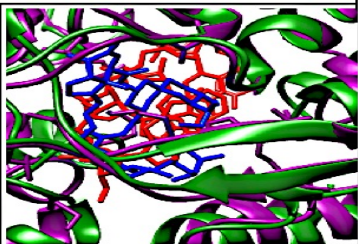
3.1 Docking result and validation

Molecular docking is one of the routinely used methods in molecular modeling and drug design. It is used to predict the conformation of small molecule (ligand) within the appropriate binding site, making it an important tool in drug discovery [46,47]. Furthermore, molecular docking ranks docked compounds based on the binding affinity of the ligand to the receptor (Figure S3).

In this study, 10 compounds were chosen to dock into the NS3 helicase based on their inhibitory characteristics at *flavivirus* ATPase/ ssRNA sites [30,31,33]. Of the 10 compounds, 3 were chosen for subsequent conformational and binding mode analysis. Lapachol and HMC-HO1 α were chosen from the naphthoquinones and purine nucleoside analogues respectively because they portrayed the most optimal docked conformation from the molecular docking studies that were carried out. Ivermectin was docked into the ssRNA site due to its high potency as a *flavivirus* inhibitor [33,48].

Validation of molecular docking was done by superimposing each of the docked complexes with the PDB structures of their natural substrates for ssRNA and ATPase site (PDB code: 5GJB and 5GJC). The results of the superimposition are shown in Table 2. (5JMT- green, 5GJB and 5GJC - magenta, ATP and RNA- red, Lapachol, HMC-HO1 α and Ivermectin –blue).

Table 7.2: 2D structure, docked complexes and validation of the docked complexes.

3D Molecular structure	Docked complex	Validation
<i>LAPACHOL</i>		
		
<i>HMC-HO1α</i>		
		
<i>IVERMECTIN</i>		
		

3.2 Molecular dynamics simulation and post molecular dynamics analysis

A frequently overlooked side of molecular docking is the flexibility of the binding target. The ligand and receptor usually undergo conformational changes before binding and sometimes the ligand fits in with little mobility. In order to ensure the stability of the complex, the 3 complexes were subjected to aMD [47,49,50].

3.2.1 Systems stability

The stability of the systems was investigated by assessing the Root Mean Square Deviation (RMSD) with regard to the C α -backbone atoms of the 3D structure during the simulation (Figure

7.2). Equilibrium was attained after 2000 ps and the overall average RMSD value for Lapachol, HMC-HO1 α and Ivermectin measured 0.98Å, 0.97Å and 0.95Å respectively. The results at the ATPase site exhibited similar stability between the HMC-HO1 α -ATPase system and the Lapachol-ATPase system, whereas, the Ivermectin-ssRNA complex demonstrated the lowest average RMSD from all three systems. This is indicative of a more stable complex, justifying Ivermectin as a potent *flavivirus* inhibitor. The RMSD plot further postulates that the binding of the three ligands at two different active sites of the protein still allowed for conformational rigidity compared to the unstable free protein, which yielded an elevated average of 1.76Å. It can therefore be deduced that all three ligands allowed for structural stability of the NS3 Helicase protein.

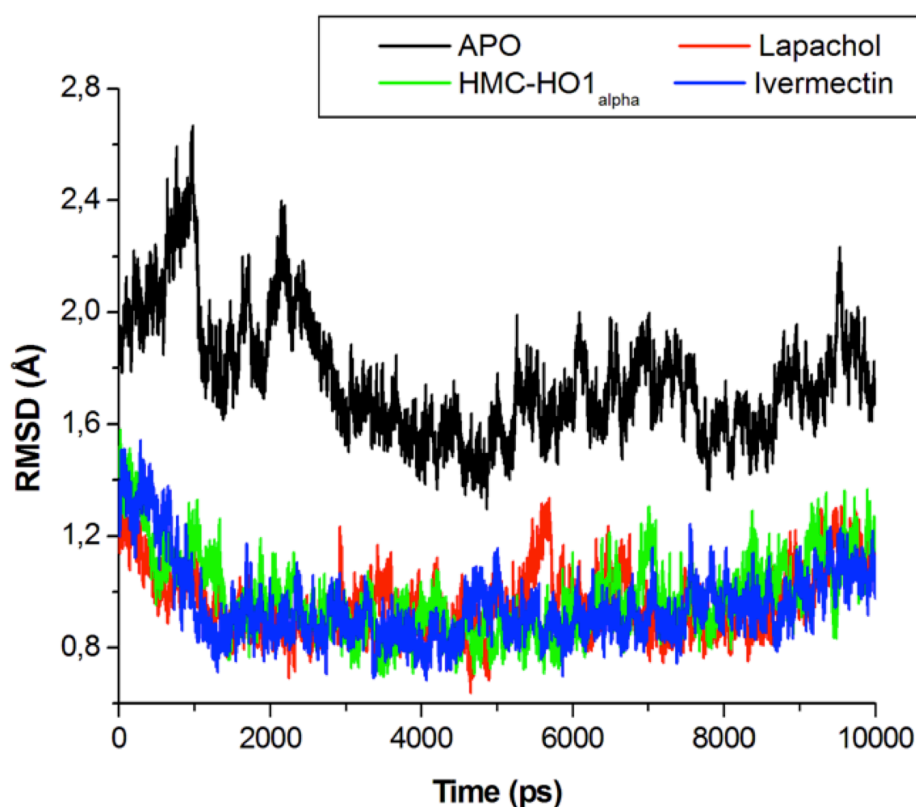


Figure 7.2: C-alpha RMSD backbone Plot for NS3 Helicase free and ligand bound conformations. The ligands Lapachol, HMC-HO1 α with Ivermectin are seen to stabilize the protein as compared to the fluctuating free protein.

3.2.2 Root Mean Square Fluctuations (RMSF)

Root Mean Square of Fluctuations (RMSF) were analyzed to show the mobility of each of the residues found in the protein, thereby giving an insight into the flexibility of the protein [51]. Figure 7.3 depicts the RMSF of the residues for each system for the duration of the simulation. High fluctuations were observed at certain residues for each of the systems, with the free protein showing the greatest fluctuations during the simulation (1.61Å). All three ligand-bound systems showed C- α residue fluctuations at residues 72-79 and 409-411. The HMC-HO1 α -ATPase system specifically showed flexibility at the “172-176” region, whereas the Ivermectin-ssRNA illustrated the lowest fluctuations of all four systems. This correlates with the RMSD stability of the systems, demonstrating the free protein to have highly unstable residues with large fluctuations compared to the ligand-bound systems. The RMSF of Ivermectin also correlated with the RMSD plot, illustrating a relatively stable system after ligand binding.

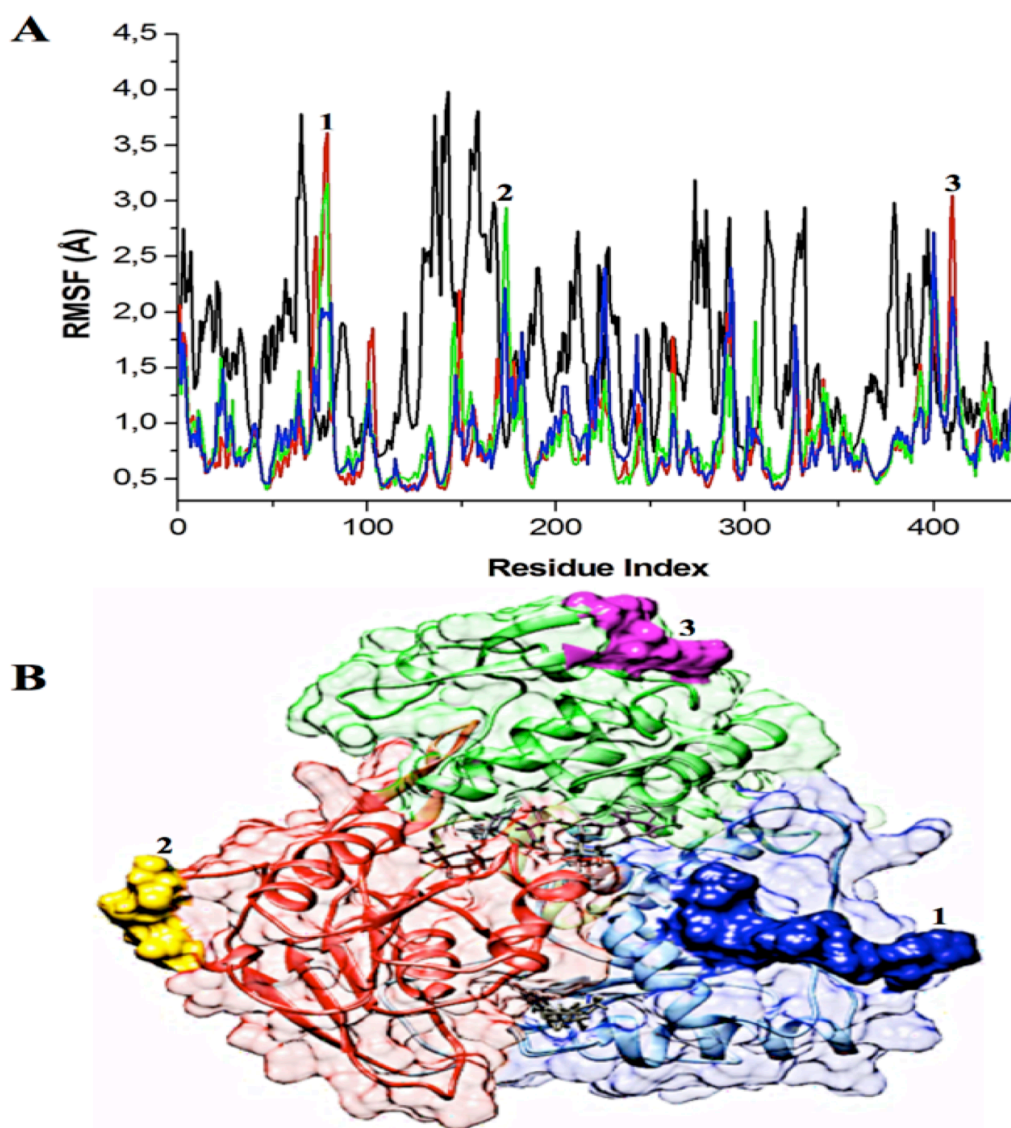


Figure 7.3: (A) RMSF Plot for Lapachol, HMC-HO1 α and Ivermectin systems. Lapachol (0.88Å) showed a higher stability at the ATPase site compared to HMC-HO1 α (0.90Å) and Ivermectin (0.85Å) showed the most favorable stability of all the systems, (B) NS3 Helicase residue fluctuations at regions: 1- the “72-79” loop (Navy), 2- the “172-176” helix (Gold) and 3- the “409-411” loop (Magenta).

3.3.3 Radius of Gyration

To further validate the stability of the systems, the overall protein shape and folding was measured by analyzing the radius of gyration (RoG) of the protein. This gave an insight into the distribution of C- α atoms within the protein [51,52]. The plots for all the systems are shown in Figure 7.4. From the graph, a difference can be seen in the compactness of the three systems from the beginning of the simulation. Ivermectin shows a lower average RoG (22.23Å) when compared to HMC-HO1 α (22.30Å) and Lapachol (22.33Å), indicating that Ivermectin exhibits a very good structural stability at the ssRNA site when it binds to the ZIKV NS3 helicase. Also at the ATPase site, the result indicates that HMC-HO1 α is more compact and therefore exhibit more stability than Lapachol. The RoG of the Apo protein correlates with the RMSD and RMSF results, showing a wide distribution of C- α atoms for the duration of the simulation, thus indicating unstable fluctuations of the protein's residues in the absence of a ligand.

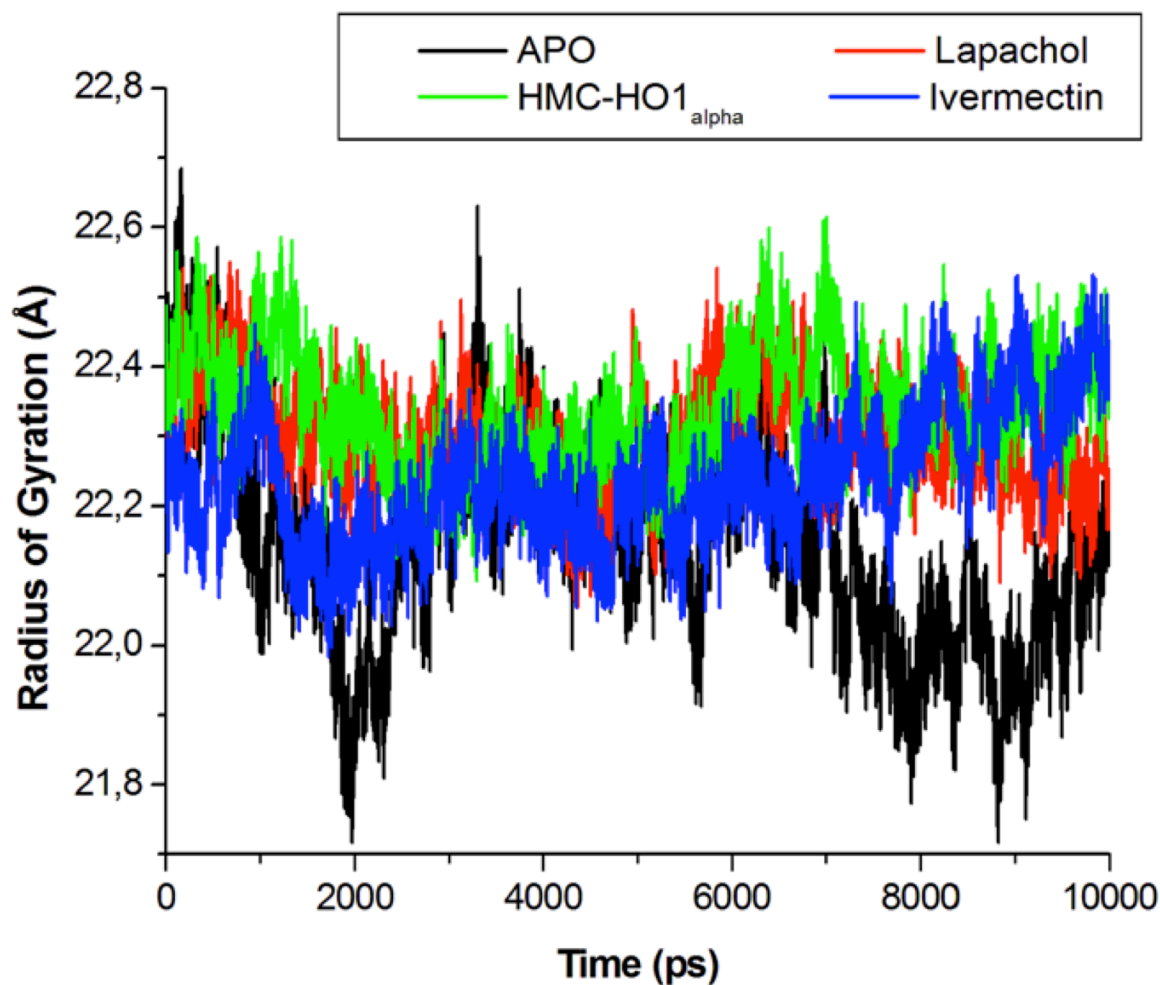


Figure 7.4: Radius of gyration Plot for Lapachol, HMC-HO1 α and Ivermectin systems when compared to the free protein.

3.3.4 Free Energy Calculations and Residue-Ligand Interaction Network

Studies have shown that free binding energies calculations are important parameters for the validation of ligand-protein binding [43]. Based on the Systems' stability, we can deduce that during the simulation, binding of the three best-docked molecules, being, Lapachol, HMC-HO1 α and Ivermectin, stabilized the fluctuating free Protein. This may be due to non-covalent interactions taking place between the ligands and the active site's residues. To estimate the binding affinities of each of the ligands to the protein, the binding free energies were calculated

using the Molecular Mechanics/Generalized-Born Surface Area method (MM/GBSA)[54]. Table 3 summarizes the binding free energy of HMC-HO1 α -ATPase and Lapachol-ATPase systems to be -42.81 kcal/mol and -39.32kcal/mol respectively. The non-polar solvation (-103.51 kcal/mol) contributed greatly towards the total binding free energy of HMC-HO1 α -NS3 helicase system while other favorable binding contributions also came from intermolecular electrostatic interactions (-62.53 kcal/mol) and van der Waals interactions (-40.98kcal/mol). Lapachol-NS3 helicase system had its greatest binding contribution from non-polar solvation energy (-65.96kcal/mol), followed by van der Waals interactions (-38.23kcal/mol) and then intermolecular electrostatic interactions (-27.73 kcal/mol). A polar solvation of 60.69 kcal/mol and 26.64 kcal/mol for HMC-HO1 α -NS3 helicase system and Lapachol-NS3 helicase system respectively were also observed. This indicates that HMC-HO1 α has a preferable binding energy than Lapachol at the ATPase site. Ivermectin had a relatively higher binding energy (-84.56 kcal/mol) at the ssRNA site with the greater energy contribution from the non-polar solvation (-136.32 kcal/mol) and van der Waals interactions (-104.36 kcal/mol).

The active site residues of proteins are important for the protein's functionality; therefore it is important to understand the interactions of these potential inhibitors with the amino acids residues in the protein [52]. In order to gain more insight into the contribution of each residue towards the binding of the ligand, per residue interaction energy decomposition analysis was carried out on the three systems.

Table 7.3: Binding free energy analysis (kcal/mol) for inhibitor- NS3 helicase complexes.

Energy Components (kcal/mol)					
Compound	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
Ivermectin	-104.36±3.95	-32.26±7.87	-136.32±11.07	52.07± 5.34	-84.56±7.77
HMC-HO1α	-40.98±3.40	-62.53±10.17	-103.51±8.78	60.69± 7.39	-42.81±4.16
Lapachol	-38.23±2.99	-27.73±6.87	-65.96±5.71	26.64± 4.12	-39.32±3.52

Lapachol-ATPase System

At the ATPase binding site, Lapachol illustrated a favorable energy contribution with residues Glu112 (-3.05kcal/mol), sharing the highest total energy, while other contributions came from residues Leu20 (-0.24 kcal/mol), Gly23 (-0.29 kcal/mol), Ala24 (-0.25 kcal/mol), Glu57 (-0.30 kcal/mol), Ala43 (-0.32 kcal/mol), Asn243 (-0.89 kcal/mol) and Arg285 (-0.8 kcal/mol). However, Lys26 (0.45 kcal/mol) and Arg288 (0.8 kcal/mol) showed unfavorable energy contributions (Figure 7.5).

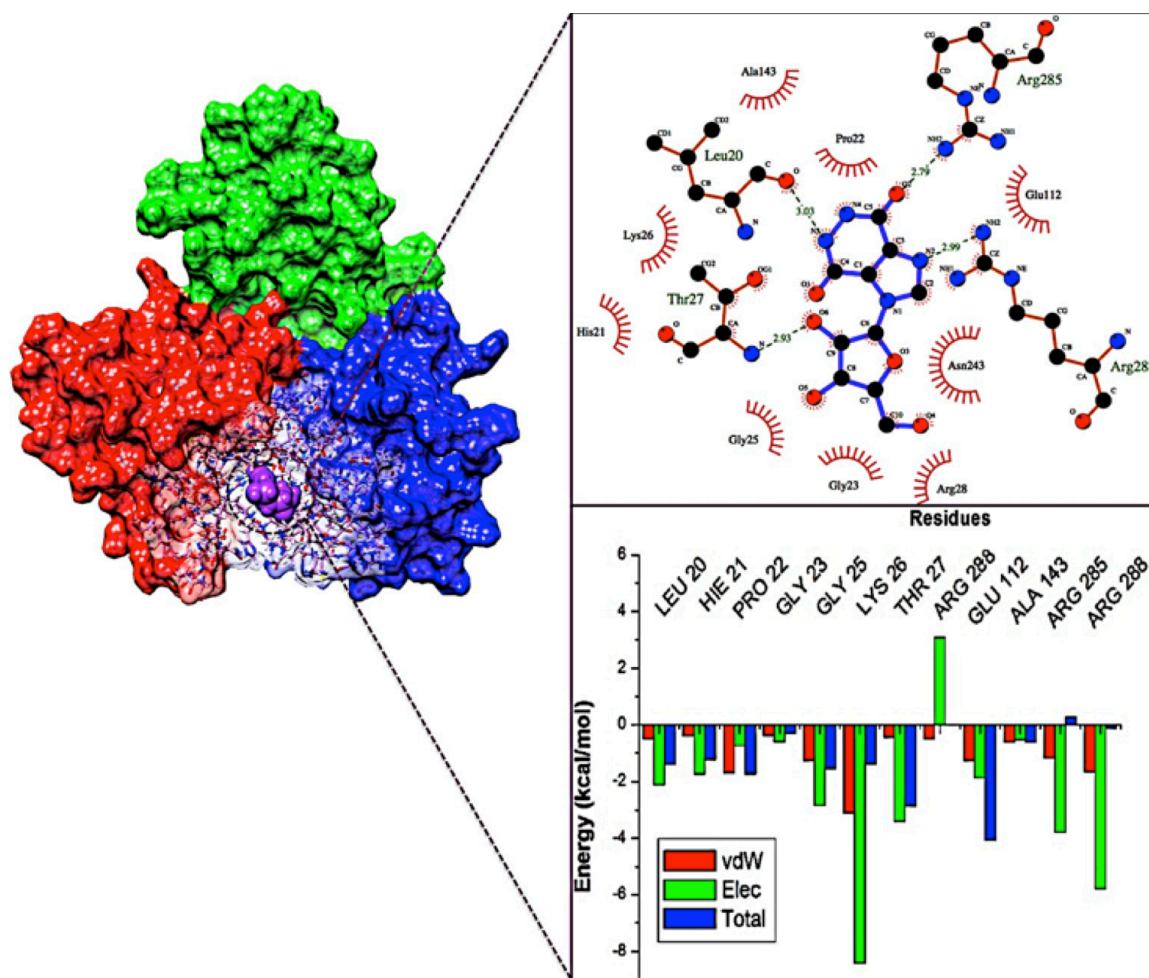


Figure 7.5: Free energy decomposition and ligand-residue interaction network at the ATPase site of the Lapachol- NS3 Helicase system.

HMC-HO1 α -ATPase System

As evident from Figure 7.5 and 7.6, HMC-HO1 α and Lapachol interact with the ATPase active sites residues by forming a hydrogen bond with residue Arg285 and hydrophobic interactions with residues His21, Gly23, Glu112 and Ala143. In addition, HMC-HO1 α exhibited hydrophobic interactions with residues Pro22, Lys26, Gly25, Arg28, and Asn243. Subsequent to HMC-HO1 α binding at the ATPase site, significant energy contributions came from residues Leu20 (-1.412kcal/mol), His21 (-1.24 kcal/mol), Pro22 (-1.75 kcal/mol), Gly25 (-1.55 kcal/mol), Lys26 (-1.40 kcal/mol), and Thr27 (-2.87 kcal/mol), with the highest contribution coming from Glu112 (-

4.08 kcal/mol), while the unfavorable energy contributions came from residue Arg285 (0.26 kcal/mol).

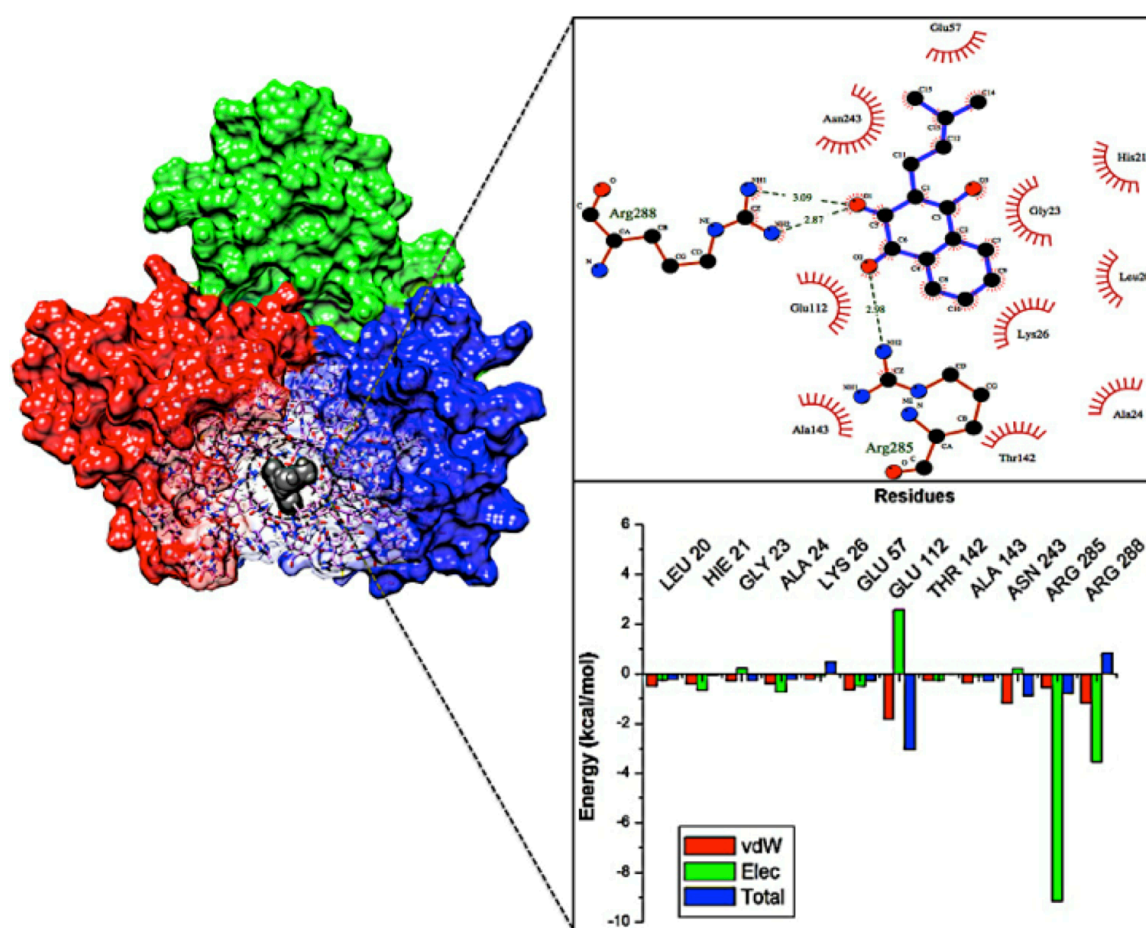


Figure 7.6: HMC-HO1α docked into the ATPase site of Zika NS3 helicase, illustrating ligand-residue interactions and active-site residue energy contributions.

The ligand-residue interaction network elucidates on the binding interactions between Ivermectin and the ssRNA active site residues, forming a hydrogen bond with Arg214 and hydrophobic

interactions with residues Ser94, Leu319, Asp117, Val369, Pro368, Thr235, Met240, Glu218, Lys215, Ala90 and Met362. The plot also reveals that residue Arg214 (-5.84 kcal/mol) had the highest total energy contribution to the binding of Ivermectin to the NS3 helicase protein at the ssRNA site. Other favorable energy contributions came from residues Ala90 (-1.48kcal/mol), Ser119 (-2.464 kcal/mol), Thr235 (-1.52 kcal/mol), Asp236 (-1.65 kcal/mol), Leu319 (-1.11 kcal/mol), Met 362 (-1.87 kcal/mol), Pro368 (-1.14 kcal/mol) and Val369 (-1.32 kcal/mol) while the unfavorable energy contribution came from Asp117 (3.0 kcal/mol).

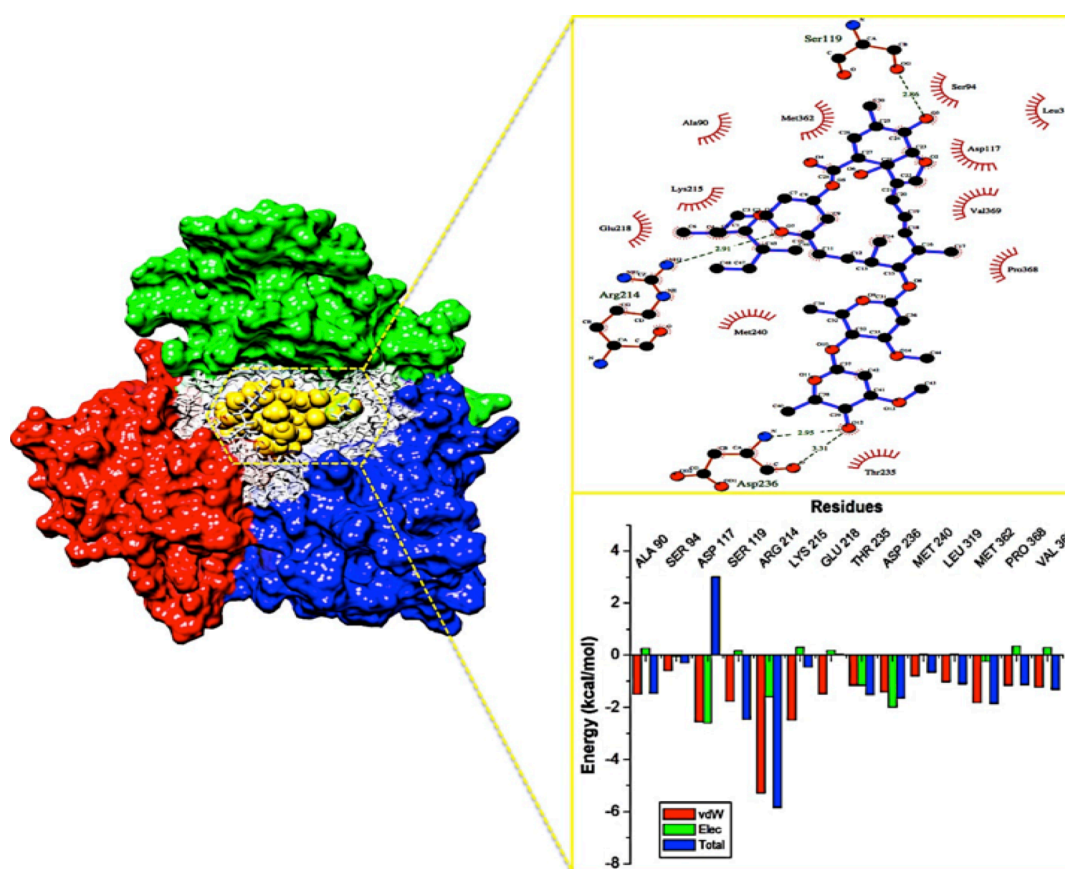


Figure 7.7: Free energy decomposition and ligand-residue interaction network at the ssRNA site of the Ivermectin- NS3 Helicase system.

4. Conclusions

In this study, we report the binding analysis of three potential inhibitors of Zika NS3 helicase at the ATPase site (Lapachol and HMC-HO1 α) and ssRNA site (Ivermectin). Results showed that

the binding of Ivermectin to ssRNA site and Lapachol and HMC-HO1 α to the ATPase site allows for conformational rigidity of the Zika NS3 helicase, thus stabilizing residue fluctuations. The interactions between the active site residues and ligands allowed for key structural flexibility at two loop regions of the NS3 helicase, thus allowing for protein stability and a possible structural mechanism of action for competitive inhibition of natural substrates.

This study aims to contribute toward the repurposing of potent *flavivirus* inhibitors against the devastating ZIKV epidemic. This strategy overcomes the concept of “shooting the dark” with experimental screening as the compounds utilized in the study have already been synthesized, thus reducing the drug discovery time-line. These potential inhibitors have been pre-clinically tested against other arboviruses and have proven to be effective [6,12–14]. Drugs such as Ivermectin have multiple functions, including anti-parasitic and more recently anti-viral properties [7,11,15]. Lapachol and HMC-HO1 α have been shown to have potent effects as *flavivirus* inhibitors including Dengue and Yellow fever virus[13].

The findings of this study provide fundamental insights toward the structural dynamics of the two active site regions on the NS3 Helicase and the ligand-receptor interaction network. The pharmacophoric characteristics found in Lapachol, HMC-HO1 α and Ivermectin may be utilized in the design of a potent hybrid drug that is able to show efficient inhibition of a multitude of diseases including the detrimental co-infection of ZIKV, Dengue and Chikungunya.

5. Future Perspective

To our knowledge, this is the first account of detailed computational investigations aimed to

provide an insight into the binding features of Lapachol, HMC-HO1 α and Ivermectin to ZIKV NS3 helicase. Based on the structural dynamics of the two active site regions on the NS3 Helicase and the ligand-receptor interaction network, it may be noted that the chemical characteristics found in these flavivirus inhibitors play a fundamental role in releasing a potent multi-purpose inhibitor against arboviruses. This will allow for pregnant females in endemic areas to take the drug as a precautionary measure against arboviruses such as Dengue and ZIKV. By having a lower toxicity and higher efficiency, a multi-purpose drug will be safe to consume by pregnant females and may diminish the risk of drug resistance due to the multiple diseases it is effective against. Distribution on a global scale and at lower cost compared to a vaccine that may need optimal storage conditions.

Author Contribution

All authors have contributed equally to this work.

Acknowledgements

The authors acknowledge the College of Health Sciences, UKZN, and the National Research Foundation for their financial support and the Center for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

References

1. Sreedharan.J. The Zika Virus: A new Threat from mosquito. *Scientia*. 11, 9–17 (2016).
2. Haddow AD, Schuh AJ, Yasuda CY, *et al*. Genetic characterization of zika virus strains: Geographic expansion of the asian lineage. *PLoS Negl. Trop. Dis*. 6(2), 1-7 (2012).
3. Martínez de Salazar P, Suy A, Sánchez-Montalvá A, Rodó C, Salvador F, Molina I. Zika fever. *Enferm. Infecc. Microbiol. Clin*. 34(4), 247–252 (2016).
4. Weaver SC, Costa F, Garcia-Blanco MA, *et al*. Zika virus: History, emergence, biology, and prospects for control. *Antiviral Res*. 130, 69–80 (2016).
5. Wikan N, Smith DR. Zika virus: History of a newly emerging arbovirus. *Lancet Infect. Dis*. 16(7), 119–126 (2016).
6. Chan JFW, Choi GKY, Yip CCY, Cheng VCC, Yuen KY. Zika fever and congenital Zika syndrome: An unexpected emerging arboviral disease. *J. Infect*. 72(5), 507–524 (2016).
7. Heang V, Yasuda CY, Sovann L, *et al*. Zika virus infection, Cambodia, 2010. *Emerg. Infect. Dis*. 18(2), 349–351 (2012).
8. Kutsuna S, Kato Y, Takasaki T, *et al*. Two cases of zika fever imported from french polynesia to Japan, December to January 2013. *Eurosurveillance*. 19(4), 1–4 (2014).
9. Hamel R, Dejarnac O, Wichit S, *et al*. Biology of Zika Virus Infection in Human Skin Cells. *J. Virol*. 89(17), 8880–96 (2015).
10. Foy BD, Kobylinski KC, Foy JLC, *et al*. Probable Non-Vector-borne Transmission of Zika Virus, Colorado, USA. *Emerg. Infect. Dis*. 17(5), 880–882 (2011).
11. Hayes EB. Zika virus outside Africa. *Emerg. Infect. Dis*. 15(9), 1347–1350 (2009).
12. Musso D, Roche C, Robin E, Nhan T, Teissier A, Cao-Lormeau VM. Potential sexual

transmission of zika virus. *Emerg. Infect. Dis.* 21(2), 359–361 (2015).

13. Besnard M, Lastère S, Teissier A, Cao-Lormeau VM, Musso D. Evidence of perinatal transmission of zika virus, French Polynesia, December 2013 and February 2014. *Eurosurveillance.* 19(13), 8–11 (2014).

14. Mahfuz M, Khan A, Mahmud H Al, *et al.* In Silico Modeling and Immunoinformatics Probing Disclose the Epitope Based Peptide Vaccine Against Zika Virus Envelope Glycoprotein. *Indian J. Pharm. Biol. Res.* 2(4), 44–57 (2014).

15. Plourde AR, Bloch EM. A Literature Review of Zika Virus. *Emerg. Infect. Dis.* 22(7), 1–15 (2016).

16. Charrel RN, Leparac-Goffart I, Pas S, de Lamballerie X KM& RC. State of knowledge on Zika virus for an adequate laboratory response. *Publ. Bull. World Heal. Organ. Type Res. emergencies Artic.*, 1–29 (2016).

17. Cao-Lormeau V-M, Blake A, Mons S, *et al.* Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet.* 387(10027), 1531–1539 (2016).

18. Baharuddin A, Hassan AA, Sheng GC, *et al.* Current approaches in antiviral drug discovery against the Flaviviridae family. *Curr. Pharm. Des.* 20(21), 3428–3444 (2014).

19. Malet H, Massé N, Selisko B, *et al.* The flavivirus polymerase as a target for drug discovery. *Antiviral Res.* 80(1), 23–35 (2008).

20. Tian H, Ji X, Yang X, *et al.* The crystal structure of Zika virus helicase: basis for antiviral drug design. *Protein Cell.* 7(6), 450–454 (2016).

21. Jain R, Coloma J, Garcia-Sastre A, Aggarwal AK. Structure of the NS3 helicase from Zika virus. *Nat. Struct. Mol. Biol.* 2(July), 1–4 (2016).

22. Incicco JJ, Gebhard LG, González-Lebrero RM, Gamarnik A V., Kaufman SB. Steady-State NTPase Activity of Dengue Virus NS3: Number of Catalytic Sites, Nucleotide Specificity and Activation by ssRNA. *PLoS One*. 8(3), 1-8 (2013).
23. Amber C. Amber 2015 Reference Manual.
24. Miao Y, Feixas F, Eun C, McCammon JA. Accelerated molecular dynamics simulations of protein folding. *J. Comput. Chem.* 36(20), 1536–49 (2015).
25. Lindert S, Bucher D, Eastman P, Pande V, Mccammon JA. Accelerated Molecular Dynamics Simulations with the AMOEBA Polarizable Force Field on Graphics Processing Units. *J. Chem. Theory Comput.* 9, 1–2 (2013).
26. Pettersen EF, Goddard TD, Huang CC, *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* 25(13), 1605–1612 (2004).
27. Windows MM V, X MOS. Molegro Molecular Viewer User Manual. (2011).
28. Da Costa ECB, Amorim R, Da Silva FC, *et al.* Synthetic 1,4-pyran naphthoquinones are potent inhibitors of Dengue virus replication. *PLoS One*. 8(12), 1–11 (2013).
29. Borowski P, Lang M, Haag A, *et al.* Characterization of Imidazo [4 , 5- d] Pyridazine Nucleosides as Modulators of Unwinding Reaction Mediated by West Nile Virus Nucleoside Triphosphatase / Helicase : Evidence for Activity on the Level of Substrate and / or Enzyme. 46(5), 1231–1239 (2002).
30. Briguglio I, Piras S, Corona P, Carta A. Inhibition of RNA Helicases of ssRNA+ Virus Belonging to Flaviviridae, Coronaviridae and Picornaviridae Families. *Int. J. Med. Chem.* 2011, 1–22 (2011).
31. Mastrangelo E, Pezzullo M, De burghgraeve T, *et al.* Ivermectin is a potent inhibitor of flavivirus replication specifically targeting NS3 helicase activity: New prospects for an old drug.

- J. Antimicrob. Chemother.* 67(8), 1884–1894 (2012).
32. Kim S, Thiessen PA, Bolton EE, *et al.* PubChem substance and compound databases. *Nucleic Acids Res.* 44(1), 1202–1213 (2016).
33. Steffen C, Thomas K, Huniar U, Hellweg A, Rubner O, Schroer A. TmoleX--A Graphical User Interface for TURBOMOLE. *J. Comput. Chem.* 31, 2967–2970 (2010).
34. Case DA, Cheatham TE, Darden T, *et al.* The Amber bimolecular simulation programs. *J. Comput. Chem.* 26(16), 1668–1688 (2005).
35. Cornell WD, Cieplak P, Bayly CI, *et al.* A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* 117(19), 5179–5197 (1995).
36. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general Amber force field. *J. Comput. Chem.* 25(9), 1157–1174 (2004).
37. Goetz AW, Williamson MJ, Xu D, Poole D, Grand SL, Walker RC. Routine microsecond molecular dynamics simulations with amber - part i: Generalized born. *J. Chem. Theory Comput.* 8, 1542–1555. (2012).
38. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79(2), 926 (1983).
39. Seifert E. OriginPro 9.1: Scientific data analysis and graphing software - Software review. *J. Chem. Inf. Model.* 54(5), 1552 (2014).
40. Greenidge PA, Kramer C, Mozziconacci JC, Wolf RM. MM/GBSA binding energy prediction on the PDBbind data set: Successes, failures, and directions for further improvement. *J. Chem. Inf. Model.* 53(1), 201–209 (2013).
41. Godschalk F, Genheden S, Söderhjelm P, Ryde U. Comparison of MM/GBSA calculations

based on explicit and implicit solvent simulations. *Phys. Chem. Chem. Phys.* 15(20), 7731–9 (2013).

42. Tsui V, Case DA. Theory and applications of the Generalized Born solvation model in macromolecular simulations. *Biopolymers*. 56(4), 275–291 (2000).

43. Kollman PA, Massova I, Reyes C, *et al.* Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* 33(12), 889–897 (2000).

44. Meng X-Y, Zhang H-X, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. *Curr. Comput. Aided. Drug Des.* 7(2), 146–57 (2011).

45. Ferreira LG, Dos Santos RN, Oliva G, Andricopulo AD. Molecular docking and structure-based drug design strategies. *Molecules*, 20(7), 13384-13421 (2015).

46. De Oliveira AS, Da Silva ML, Flávia A, *et al.* NS3 and NS5 Proteins: Important Targets for Anti-Dengue Drug Design. *J. Braz. Chem. Soc.* 25(10), 1759–1769 (2014).

47. Jr. FRS. Molecular Dynamics Simulations of Protein Dynamics and their relevance to drug discovery. *Pharmacol.* 10(6), 738–744 (2010).

48. Arodola OA, Soliman MES. Could the FDA-approved anti-HIV PR inhibitors be promising anticancer agents? An answer from enhanced docking approach and molecular dynamics analyses. *Drug Des. Devel. Ther.* 9, 6055–6065 (2015).

49. Sing A SM. Understanding the cross-resistance of oseltamivir to H1N1 and H5N1 influenza A neuraminidase mutation using multidimensional computational analyses. , 1–18 (2015).

50. Appiah-Kubi P, Soliman MES. Dual anti-inflammatory and selective inhibition mechanism of leukotriene A4 hydrolase/aminopeptidase: insights from comparative molecular dynamics and binding free energy analyses. *J. Biomol. Struct. Dyn.* 1102, 1–16 (2016).

51. Karubiu W, Bhakat S, Soliman MES. Compensatory role of double mutation N348I/M184V on nevirapine binding landscape: insight from molecular dynamics simulation. *Protein J.* 33(5), 432–446 (2014).
52. Kumalo HM, Soliman ME. Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced In Silico Approach in Drug Discovery: A Case Study on the Identification of Novel Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents. *Cell. Mol. Bioeng.* 9(1), 175–189 (2016).

CHAPTER 8

8.1 Conclusion

Since the outbreak of the epidemic in 2015, ZIKV has been evidenced to manifest in a calamitous manner. One of the most devastating effects of the disease is the onset of ZIKV-related microcephaly in neonates primarily through prenatal transmission from the infected mother. Finding a preventative cure or treatment regimen against this virus has thus become paramount. During 2016, a flood of previously unknown information regarding the disease was released, with researchers worldwide working tireless in the design of potential vaccines or small drug molecules as potential ZIKV inhibitors.

This study sought to augment the structural and molecular characteristics of this previously neglected tropical disease by identifying the principal target proteins of ZIKV, being the NS3 Helicase, NS5 MTase and NS5 RdRp, creating CADD route map to identify potential ZIKV inhibitors and finally, implementation of the route map to investigate potential inhibitors for all three viral targets.

Upon commencement of the study, no crystal structures of ZIKV target proteins were available; therefore a homology modeling was implemented to generate a 3D model of the NS5 protein, proving to be an invaluable computational tool. Thereafter, the NS5 MTase and RdRp were subjected to “per-residue energy decomposition pharmacophore” virtual screening to identify favorable molecules that are more effective and less toxic than the experimentally tested inhibitors. The most favorable inhibitors identified in this study for NS5 MTase and RdRp are BG323 and ZINC39563464, respectively.

The NS3 Helicase was also investigated in this study as it plays a major role in viral replication. In this domain of the study, the structural characteristics of the ATPase active site were explored when bound to a competitive inhibitor. A clear ligand-dependent flexibility was noted in the active site/P-loop and a stabilizing α -Helix formed after ligand binding. These insights into the binding landscape will aid other researchers in finding a potent, yet effective inhibitor against ZIKV. To round up this study, potent *Flavivirus* inhibitors were selected based on ZIKV’s structural similarity to other *Flaviviruses*, including Dengue and West Nile (Jain *et al.* 2016; Tian

et al. 2016). These inhibitors were then assessed, based on their free binding energies at their respective ATPase or ssRNA active; HMC-HO1 α and Ivermectin were reported as the most favorable compounds.

Overall, this study has provided valuable insights into the design and development of ZIKV inhibitors through molecular modeling and CADD.

8.2 Future Perspectives

The potential inhibitors of the study have presented promising protein-ligand interactions and binding energies and therefore maybe utilized as the lead compounds. However, prospective biological testing of these compounds is still required to verify these *in silico* studies.

The ramifications of ZIKV infection have led to a multitude of potential small molecule inhibitors and at least two vaccines that are currently in clinical trials. However, the virus has other challenging defense mechanisms that may render most drugs ineffective. Zika is able to penetrate the blood-brain-barrier (BBB) as evidenced by its downstream pathological effects on the nervous system (Huang *et al.* 2016; White *et al.* 2016; Anaya *et al.* 2016; Barrows *et al.* 2016). The BBB however, only allows entry to hydrophilic drugs due to their tight junctions that form as a point of entry. If drugs are not able to pass the BBB, the ability of the drug to inhibit the target protein in neuronal cells becomes unsuccessful. In order to overcome this obstruction, two approaches may be used:

1. The design (*in silico*) and synthesis of potential compounds that adhere to the requirements for permissivity.
2. The use of a drug delivery system to allow for the drug to pass through the BBB. This system may be in the form of a nanoparticle, polymer or aptamer.

With regard to this investigative study, a purely computational perspective was used. However, to further the ZIKV drug design research toward targeted therapy, an experiment will need to be designed to analyze the efficacy of the two approaches mentioned above.

References:

- Anaya, J.-M. *et al.*, 2016. Zika virus and neurologic autoimmunity: the putative role of gangliosides. *BMC Medicine*, 14(1), pp.1–3.
- Barrows, N.J. *et al.*, 2016. A Screen of FDA-Approved Drugs for Inhibitors of Zika Virus Infection. *Cell Host and Microbe*, 20(2), pp.259–270.
- Huang, W.-C. *et al.*, 2016. Zika virus infection during the period of maximal brain growth causes microcephaly and corticospinal neuron apoptosis in wild type mice. *Scientific Reports*, 6(34793), pp.1–8.
- Jain, R. *et al.*, 2016. Structure of the NS3 helicase from Zika virus. *Nature Structural & Molecular Biology*, 23(8), pp.752–754.
- Tian, H. *et al.*, 2016. The crystal structure of Zika virus helicase: basis for antiviral drug design. *Protein & Cell*, 7(6), pp.450–454.
- White, M.K. *et al.*, 2016. Zika virus: An emergent neuropathological agent. *Annals of Neurology*, 80(4), pp.479–489.

APPENDIX

Appendix A:

Pritika Ramharack and Mahmoud E.S. Soliman (2016) Zika Virus Drug Targets: A Missing Link In Drug Design And Discovery – A Route Map To Fill The Gap, *RSC Advances*, 6, 68719-68731.

Appendix B:

Pritika Ramharack and Mahmoud E.S. Soliman (2017) Zika Virus NS5 Protein Potential Inhibitors: An Enhanced *in silico* Approach in Drug Discovery, *Journal of Biomolecular Structure & Dynamics*, 2017, 1-16.

Appendix C:

Pritika Ramharack, Sofiat Oguntade, Mahmoud E.S. Soliman (2017) Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery, *RSC Advances*, 7, 22133-22144.

REVIEW

Cite this: *RSC Adv.*, 2016, 6, 68719

Zika virus drug targets: a missing link in drug design and discovery – a route map to fill the gap†

Pritika Ramharack and Mahmoud E. S. Soliman*

Zika virus is an emerging virus that has been defined by the World Health Organization as a serious global biological-threat. Zika virus is an arbovirus from the *flavivirus* genus that is linked to microcephaly after prenatal transmission from the infected mother and most recently Guillain-Barré Syndrome. The need for innovative research methods is urgent due to the ambiguity surrounding Zika virus. The lack of experimental data regarding potential drug targets, strategies for design and drug resistance has prompted us to provide a comprehensive framework with structured theoretical and technical guidelines on potential drug targets, modeling and design of inhibitors against the virus, thus assisting and encouraging scientists from different research domains to fill the gap in this research area. We have also presented a 3D homology model of the ideal Zika viral target, the non-structural protein 5, identified the active binding sites of each domain of the protein and found potential compounds that may act as inhibitors. This report will be immensely beneficial toward the design of Zika virus drug inhibitors.

Received 10th May 2016

Accepted 11th July 2016

DOI: 10.1039/c6ra12142j

www.rsc.org/advances

1. Introduction

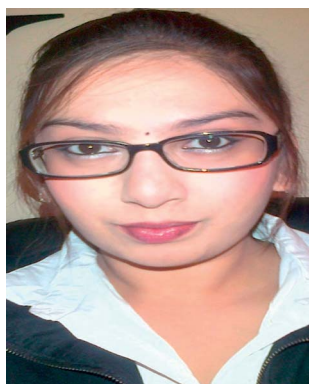
Zika virus (ZIKV) is a re-emerging arthropod-borne virus that is predominantly found in the tropics. However, rapidly evolving climate conditions coupled with increasing distribution of *Aedes* mosquito vectors and emerging modes of transmission of the virus have increased the potential to cause outbreaks in previously unaffected areas.¹ The virus is a member of the Spondweni serocomplex of the genus *flavivirus*, family

flaviviridae. Other arboviruses related to ZIKV include Dengue virus, Japanese encephalitis viruses and West Nile virus.^{2,3}

The first cases of ZIKV infection were reported in Nigeria in the 1950's. Since then, ZIKV has shown erratic cases in countries such as Uganda, Tanzania, Egypt, Gabon, and in parts of Asia including India and Indonesia, with the most devastating pandemic occurring in Brazil in 2015.^{4,5} Since the outbreak in Brazil, infection has spread rapidly throughout South America and Mexico, with Colombia being one of the most-affected countries with over 20 000 suspected cases.⁶ As of June 2016,

Molecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa. E-mail: soliman@ukzn.ac.za; Fax: +27 31 260 7872; Tel: +27 31 260 8048

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c6ra12142j



Pritika Ramharack is a PhD student under the supervision of Prof M.E. Soliman at the Department of Pharmaceutical Sciences, University of Kwa-Zulu Natal, South Africa. Her current research concentrates on the Design of Inhibitors against the Zika virus using in silico studies.



Prof. Mahmoud Soliman (B. Pharm., M. Pharm., MPhil/PhD) is the Dean and Head of School of Health Sciences, University of Kwa-Zulu Natal, South Africa, and the Head and principal investigator of the Molecular Modeling and Drug Design Laboratory. Prof. Soliman is also an academic visitor at the University of Bath, UK, collaborating with Prof. Ian Williams' lab. Soliman's

research mainly focuses on studying biomolecular systems and drug-protein interactions at a molecular level.

thirteen countries have reported Central Nervous System (CNS) malformations such as microcephaly and Guillain-Barré syndrome (GBS) which may potentially be linked to ZIKV; during the recent circulation of the virus, eight countries had reported cases of GBS, where laboratory testing confirmed ZIKV infection in a number of those cases.⁷ Globally, the prevalence of ZIKV infection may be greatly underestimated (Fig. 1) due to the recently verified prenatal and sexual transmission in humans,⁸ as well as the abstruseness surrounding the pathogenicity and thus, in turn the search of inhibitors of this “neglected disease”.

Most cases of ZIKV have reported febrile flu-like conditions that may be mistaken for other viral infections such as yellow fever. Other symptoms include swollen lymph nodes, maculopapular skin rashes and joint pains.^{5,9,10} Current research has raised concerns that the virus could cause dramatic increases in microcephaly in newborns after prenatal transmission.^{6,11–13} Complications associated with prenatal infection encompass fetal growth restriction, neurological and ocular abnormalities, intracranial calcification and in some cases perinatal death or stillbirth.^{10,14}

The virus is transmitted *via* an *Aedes* mosquito vector, congenital and perinatal transmission, as well as sexual intercourse.^{8,11,14–17} Studies have also reported transmission *via* blood transfusion and laboratory exposure.^{8,18} Commercial assays have been utilized in the diagnosis of ZIKV infection, including Real Time Polymerase Chain Reaction (RT-PCR) kits and IgM-based enzyme-linked immunosorbent assay (ELISA).^{10,19} Sample DNA and RNA for these kits may be extracted from blood serum, semen, amniotic fluid, plasma, saliva and urine. In Dengue infection, NS1 protein may be detected in a host's blood serum at the onset of clinical

symptoms, this may prove to be another method by which ZIKV infection may be identified.^{16,20–22}

Although recent publications have described the global spread, pathogenicity and bioinformatics of ZIKV and its comparison between other flaviviruses including Dengue, West Nile, Yellow fever and Japanese Encephalitis virus,^{1,3,8,10,13,19,20,23–30} fundamental research into ZIKV small molecule drug design will be key in developing inhibitors of target proteins of the virus. Ekins *et al.* (2016) described possible drug discovery and potential homology models of multiple proteins of ZIKV, however, despite the execution of research methods, there are currently no known FDA approved drugs of ZIKV.³¹ This prompted us to conduct a concise route map depicting the steps taken toward identifying potential inhibitors of drug targets with no 3D crystal structure and by following the guide to create a homology model of a non-structural protein of the virus, thus assisting scientists from different research domains. These *in silico* guidelines will be vastly beneficial in aiding and accelerating ZIKV experimental drug discovery.

2. Overview of ZIKV protein assembly

ZIKV is an enveloped virus comprising of an 11 kilobase, single-stranded positive sense RNA genome consisting of 10,794 nucleotides encoding 3,419 amino acids.²⁵ The open reading frame (ORF) of the 5' and 3' untranslated region (UTR) encodes a polyprotein that is cleaved into three structural proteins being the capsid, precursor membrane, and envelope. Seven non-structural (NS) proteins are also found in this assembly, namely, NS1, NS2A, NS2B, NS3, NS4A, 2K, NS4B, and NS5 (largest viral protein).²⁴ These viral assembly proteins may act as crucial molecules in drug discovery.

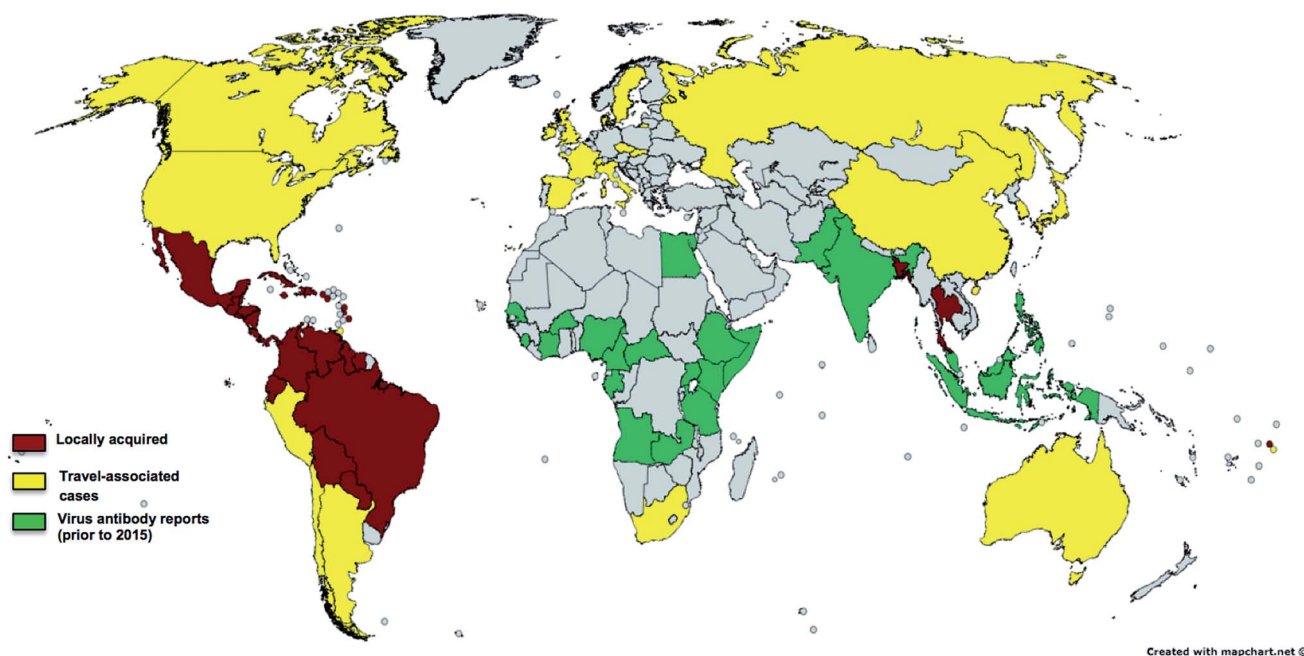


Fig. 1 Global reports of ZIKV transmission, infection and sporadic viral antibody reports prior to 2015, as of April 2016 (adapted from Centers for Disease Control and Prevention).

3. Potential biological drug targets against ZIKV

3.1 Viral drug targets

Hughes *et al.* (2010) stated that the potential of a protein as a therapeutic target and its effectiveness in drug design is essential for determining the biological utility of the protein.³² ZIKV contains viral proteins that may act as targets in drug design (Table 1).

The structural proteins of ZIKV, being the capsid, precursor membrane and envelope form the viral particle.³³ The envelope (E) protein is the key surface protein as it is able to mediate various aspects, including binding and membrane fusion of the viral replication cycle, making it a significant target in drug design.⁵

The nonstructural proteins participate in the replication of the RNA genome, virion assembly and invasion of the innate immune system. Of the nonstructural proteins, NS5, NS3 and NS1 have shown enzyme activity in other viruses of the *flavivirus* genus, creating ideal targets in inhibitor development.³³

NS5 is a bifunctional enzyme with a methyltransferase domain at its N-terminal end and a RNA-dependent RNA polymerase (RdRp) at its C-terminal end. Both N- and C-terminal domains contain an *S*-adenosyl-Methionine-dependent MTase core structure that folds into an $\alpha/\beta/\alpha$ sheet cradled between the N- and C-terminal subdomains.³⁴ The protein engages in virus–host interactions and actively interacts with the host environment.¹ To our knowledge, there is currently no available 3D crystal structure of the ZIKV NS5 protein.

The NS3 protein is a multifunctional, viral replication protein. The protease comprises of the N-terminal third of NS3 and nucleotide triphosphatase, the RNA triphosphatase, and finally the helicase components. NS3 can be considered a serine protease and contains a classical catalytic triad (His-51, Asp-75, Ser-135).^{1,35} Agnihotri *et al.* (2012) reported an *in silico* study in which a homology model of the *flavivirus* NS3 protein was created using 22 species of the *flavivirus* genus. This study is

a critical tool in the understanding the *flavivirus* NS3 protein and thus the impact of the protein as a ZIKV target.³⁶ The 3D-crystal structure of the NS3 helicase protein has recently been reported in *Protein and Cell* where a conserved triphosphate pocket and a positively-charged tunnel were identified to be critical for the hydrolysis of nucleoside triphosphates and the accommodation of RNA respectively.³⁷

The 3D crystal structure of the noteworthy NS1 glycoprotein viral target was released earlier this year and was classified as a major antigenic marker of ZIKV infection.³⁸ The NS1 is synthesized as a monomer and dimerizes after post-translation modification in the replication cycle.³⁹ The mature NS1 protein has significant immune evasive functions on the surface of cells, in the extracellular space and in cells by directly regulating the translation of viral RNA. Recent studies on Dengue virus have also evidenced NS1 to be associated with vascular leak and shock due to the disruption of TLR3 signaling pathways.⁴⁰ Song *et al.* (2016) reported NS1 to display a loop-surface interface with divergent electrostatic potential that may result in unique interactions with host machinery compared to that of other flaviviruses.³⁸ This makes ZIKV NS1 an ideal target for chemoinformatics studies.

Inhibitors of these viral proteins may be designed using computer-aided drug design techniques to select structural molecules that may inhibit the replication of viruses such as ZIKV in a host.

3.2 ZIKV host targets

During the ZIKV replication cycle, host cell machinery is imperative in the translation of viral RNA and maturation of the replicated virus, thus targeting host proteins and pathways may be key to effective inhibition of viral replication.

One of the most researched host proteins in *flavivirus* infection is the endoplasmic reticulum glucosidase. These proteins allow for the cleavage of the terminal glucose from the glycan found at the glycosylation-site of the prM and envelope protein, thus leading to its maturation of the envelope protein.⁴¹

Table 1 Potential ZIKV target proteins^a

Protein	NCBI reference sequence	PDB code	Residue count
Structural proteins			
Capsid	YP_009227206.1	5IZ7/5IRE	122aa
Precursor membrane	YP_009227197.1	5IZ7/5IRE	168aa
Envelope	YP_009227198.1	5JHM/5JHL	500aa
Nonstructural proteins			
NS1	YP_009227199.1	5IY3	352aa
NS2A	YP_009227200.1	Not available	226aa
NS2B	YP_009227201.1	Not available	130aa
NS3	YP_009227202.1	5JMT	617aa
NS4A	YP_009227203.1	Not available	127aa
2K	YP_009227209.1	Not available	23aa
NS4B	YP_009227204.1	Not available	251aa
NS5	YP_009227205.1	Not available	903aa

^a aa-amino acid.

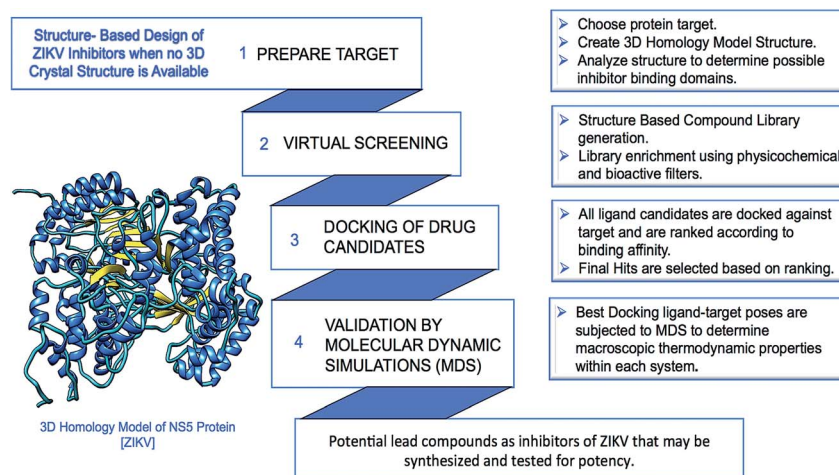


Fig. 2 Route map toward the *in silico* design of ZIKV inhibitors using the homology modeled viral NS5 protein.

Studies have shown that many flaviviruses, including ZIKV, have a *N*-glycosylation at Asn154.^{31,42} Castanospermine (CST) and deoxynojirimycin (DNJ) have been established as potent inhibitors of alpha-glucosidases, thus preventing the early stages of glycosylation.⁴³

Hamel *et al.* (2015) described the importance of dendritic cell-specific intracellular adhesion molecule 3-grabbing non-integrin (DC-SIGN), TIM and TAM receptors in the attachment and entry of ZIKV into the host cell before replication can occur.⁴⁴ Small interfering RNA (siRNA) was also shown to completely inhibit the expression of the above proteins after 48

hours. Other informative publications on siRNA inhibition of *flavivirus* host machinery include a review by Hirsch (2010), an *in silico* based experimental study on Dengue virus by Noppakunmongkolchai *et al.* (2016) and the silencing of the 3' UTR of ZIKV genome by Shawan *et al.* (2015).^{5,45,46}

A recent study published in June by *Nature* identified host endoplasmic reticulum-associated signal peptidase complex (SPCS) to be necessary for the proper cleavage of ZIKV prM and envelope proteins. The authors also demonstrated that the loss of SPCS signaling leads to a dramatic decrease in Dengue, Yellow fever, West Nile, JEV and Hepatitis C viruses. This study

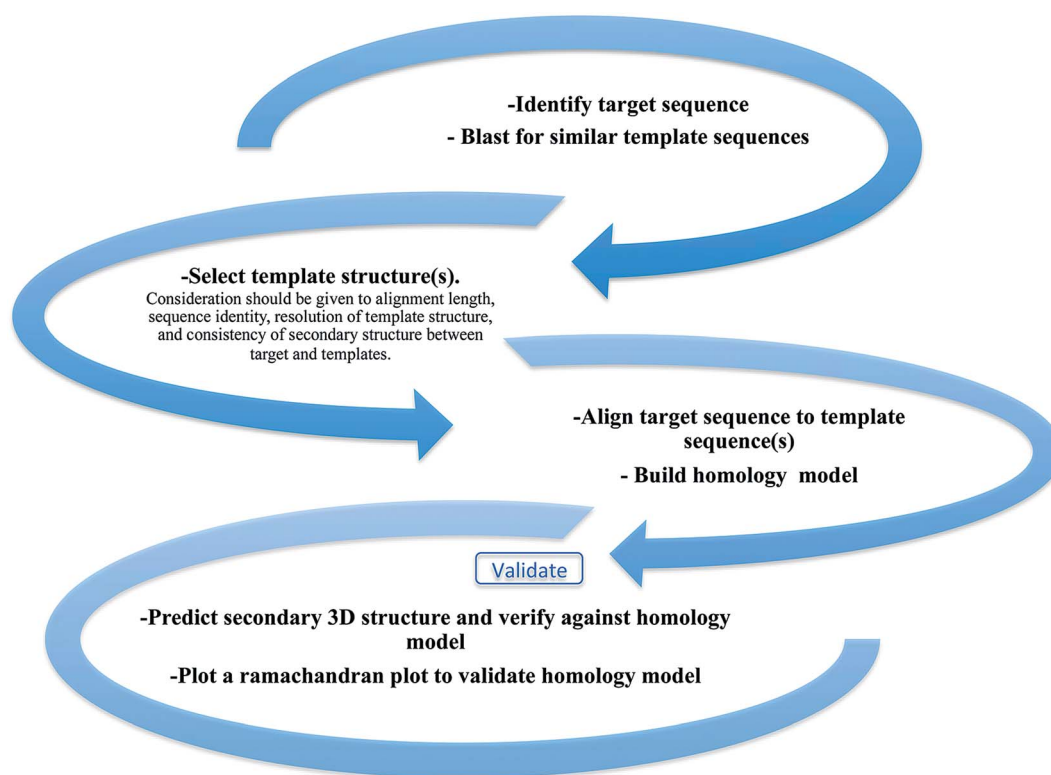


Fig. 3 Protocol for building a homology model in our laboratory.

could be a critical cornerstone in targeting host proteins and pathways in ZIKV infection.⁴⁷

4. *In silico* studies conducted on ZIKV

Prior to 2016, only two *in silico* reports have been made toward the development of ZIKV inhibitors. Computational studies by Shawan *et al.* (2014) showed the viral envelope glycoprotein to be the most immunogenic structural protein of the virus, thus, making it a candidate for vaccine development.⁵ Mahfuz *et al.* (2015) also looked at small interfering RNA (siRNA) in gene silencing of the 3' UTR of ZIKV genome.⁴⁸ Following the Brazil outbreak, an influx of research output has flooded the scientific community. There has been numerous computational studies regarding ZIKV target proteins; crystals structures of the NS1, NS3, envelope and the 2 cryo-EM structures of the stable virus have been released.^{37,38,42,49,50} Ekins *et al.* (2016) described *in silico* studies in both drug discovery and the homology models of both structural and nonstructural proteins.^{9,31} There have also been reports comparing the structural and sequence conformations of ZIKV to other flaviviruses including Dengue and West Nile viruses.^{51,52}

5. *In silico* route map toward the design and discovery of ZIKV inhibitors

Rational drug design may be classified into two groups, the first being the development of small molecules with the desired effects of the target, whose structural information is known and the second group being development of small molecules whose cell functions and structural information may not be known.⁵³

To date, there is no available 3D crystal structure of the ZIKV NS5 protein. This prompted us to create a route map (Fig. 2) describing the techniques of the second group, thereby benefiting scientists from different research domains by informing them of fundamental computational techniques in the design of novel small drug molecules, allowing for increased output of validity experimentation.

Computer-Aided Drug Design (CADD) represents computational methods and resources used in the design and discovery of new therapeutic solutions.⁵⁴ Numerous bioinformatics tools and resources have been developed to advance the drug discovery process.^{55,56} The recent improvements made in computational chemistry software, CADD and molecular dynamic simulations have led to innovative research methods in the pharmaceutical industry.⁵⁷

Details on how the homology model was created are described under Section 5.1.

The initial step of any modeling work is having a valid 3D structure, from X-ray crystallography, Nuclear Magnetic Resonance (NMR) or computational design using homology modeling. Homology modeling is used to predict and generate a plausible 3D structure of ZIKV's biological target from a template sequence based on the structure of one or more homologous viral proteins of which crystal assemblies have been reported (Fig. 3).⁵⁸

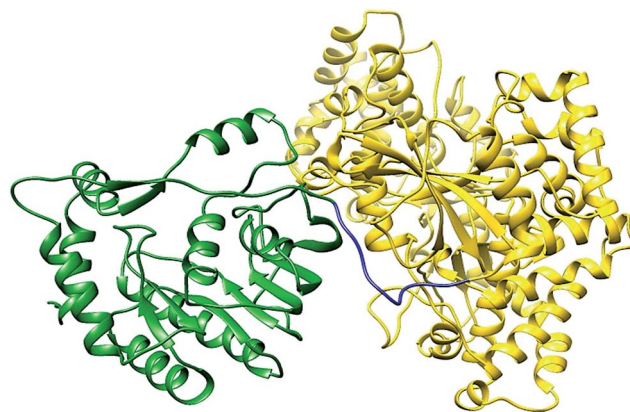


Fig. 4 Homology model of ZIKV NS5 protein.

In order to create a ZIKV target 3D homology model, a typical procedure needs to be followed, with validation taking place at each step. Once the 3D structure has been generated and validated using 3D profiles and a Ramachandran plot,⁵⁹ the predicted active binding site of the ZIKV target molecule may be identified. If the drug target is an enzyme, such as the NS3 or NS5 protein of the ZIKV viral assembly, designed chemical molecules may be able to fit within an active site pocket.⁵⁶ The results establish the locality of possible binding pockets of the protein.⁶⁰ After each pocket has been identified, we can identify the size of a pocket (volume, surface area and depth), possible interacting residues and surface atoms.⁶¹

Subsequent to the ZIKV protein homology model and target site determination, several paths may be utilized in the

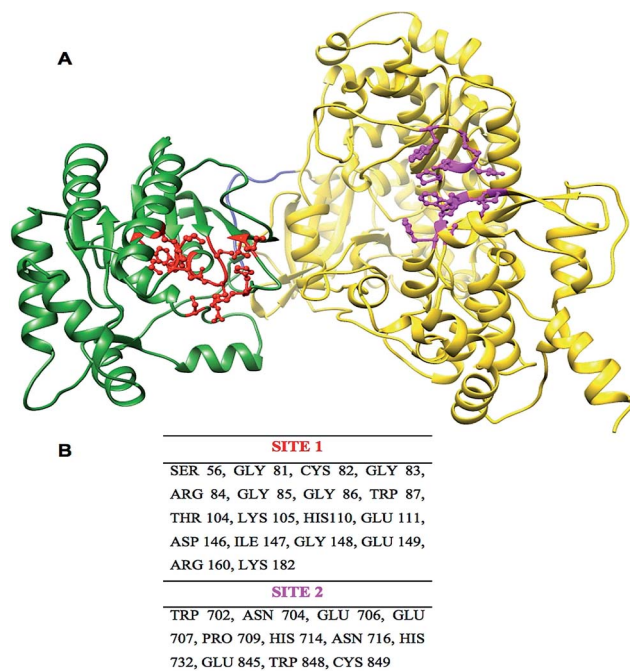


Fig. 5 The potential binding sites, identified by Site-hound,⁷¹ of ZIKV NS5 protein. (A) Site 1 (methyltransferase active binding site) (red) and Site 2 (RdRp active binding site) (magenta), (B) active binding site residues of the NS5 protein at Site 1 and Site 2.

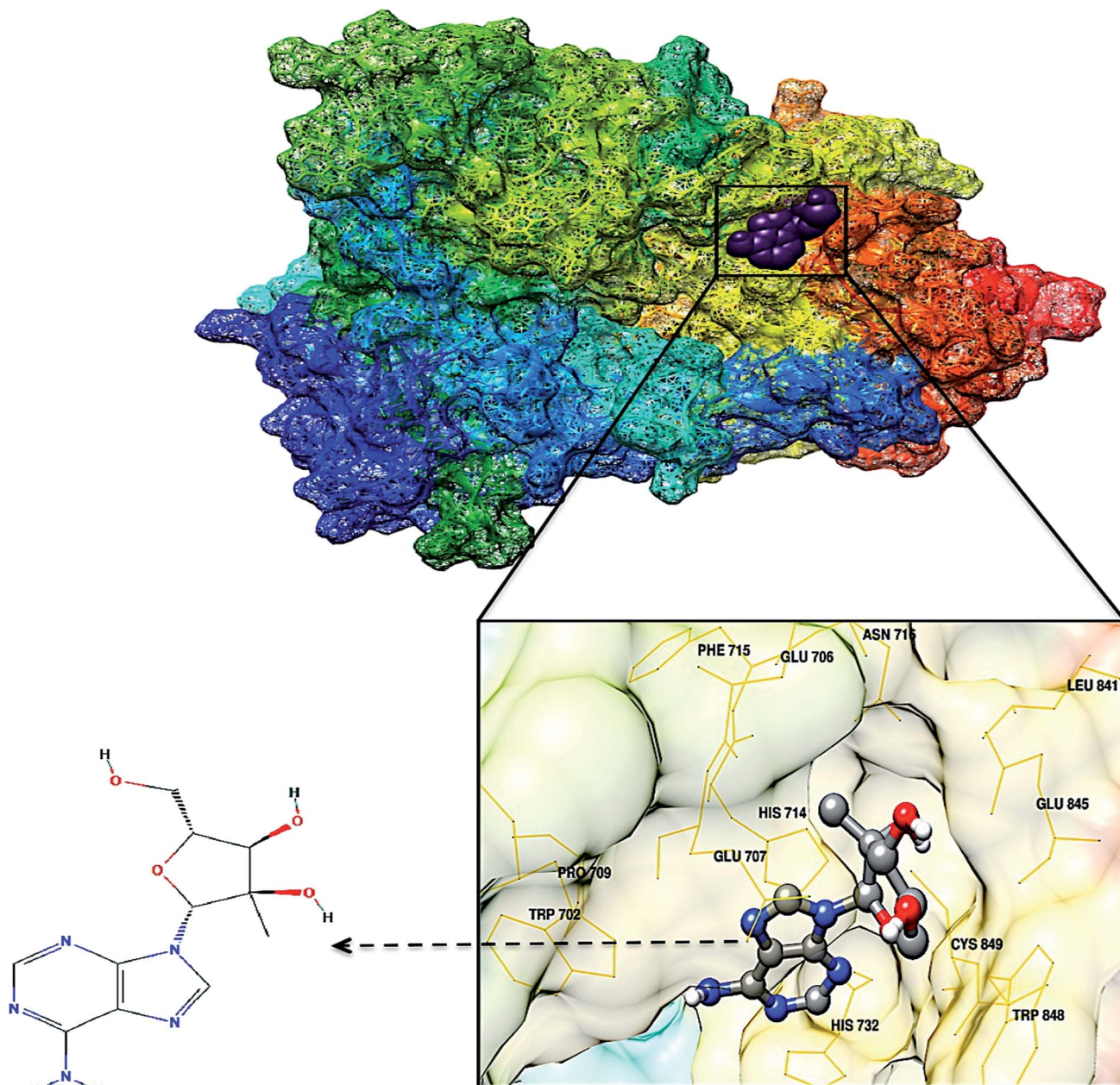


Fig. 6 Docked conformation of 2'-C-methyladenosine with ZIKV NS5 RdRp (binding affinity: $-6.3 \text{ kcal mol}^{-1}$).

development of inhibitors. Structure-based virtual screening will assist in searching through combinatorial chemistry libraries for molecules that may be potential inhibitors of the target protein and automatically dock these molecules into the 3D target's active pocket at a rapid rate.⁶² Thousands of molecules may be able to match the active site of the target protein, thus, a scoring function is utilized to rank ligands based on the free binding energy calculated after each docking pose.^{63,64} Molecules with the lowest free binding energy subsequent to screening may be used as inhibitor candidates, which may then be employed in a series of validatory molecular dynamic simulations.

Molecular dynamic simulations calculate the trajectory of a generated docking pose by utilizing Newtonian mechanics.⁶⁵

It is an important tool of CADD as it avoids analytic intractability in complex systems.⁵⁷ Molecular dynamics is not essential in CADD but it can provide validation of docking results between a protein and its potential inhibitors.⁶⁶

By implementing *in silico* studies in the design of ZIKV protein inhibitors, putative drug-like compounds may be identified and their potency verified using *in vitro* and *in vivo* testing.

Studies report that *in vitro* testing of potential inhibitors may utilize cultured monkey cell lines such as LLC-MK2 and Vero.³⁰ Delvichio *et al.* (2016) also reported chloroquine as potential ZIKV inhibitor in Vero, hBMEC, hNSC and mouse neurospheres. Dowall *et al.* (2016) developed the first *in vivo* murine

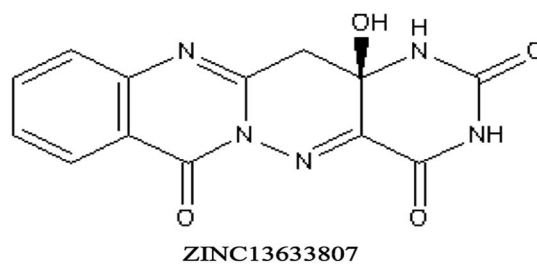
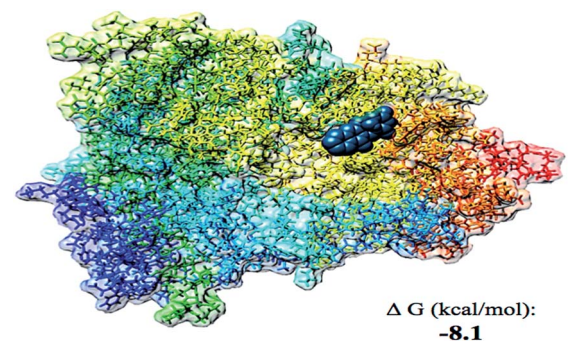
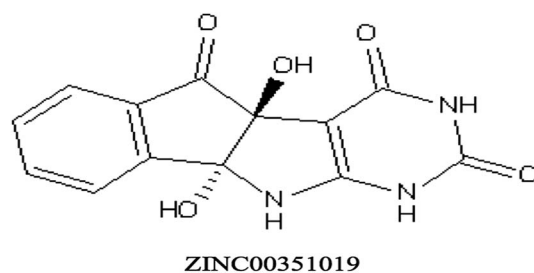
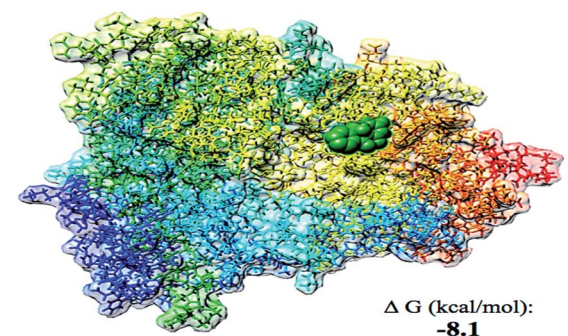
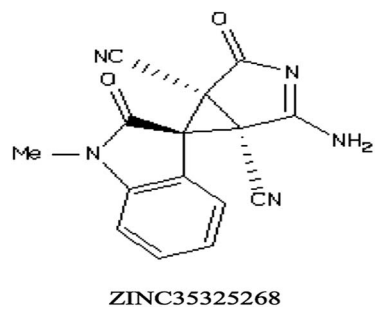
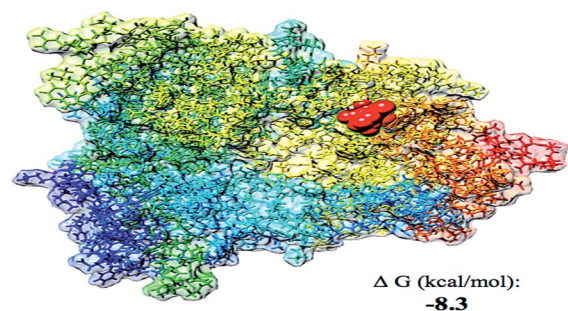
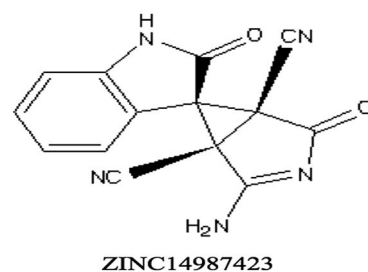
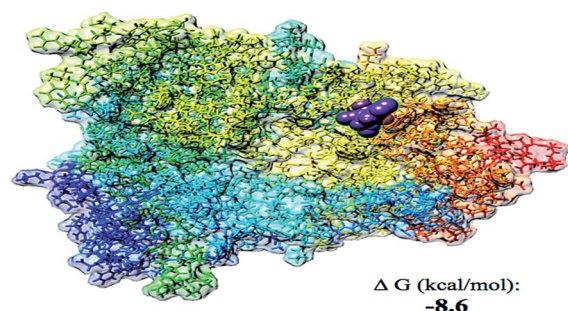
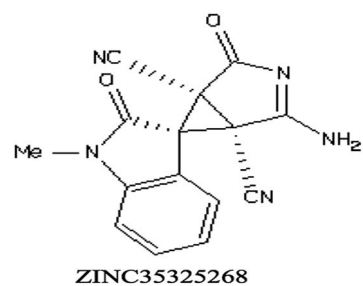
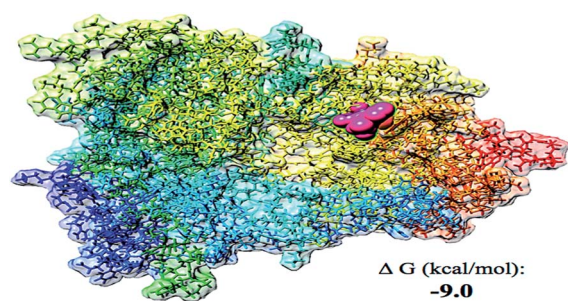
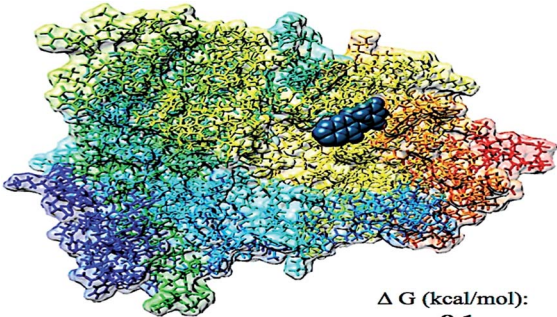
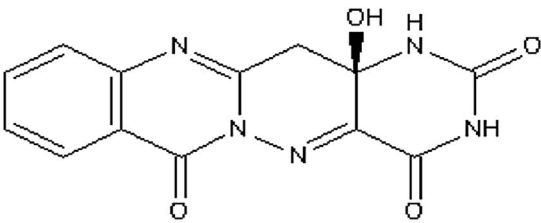
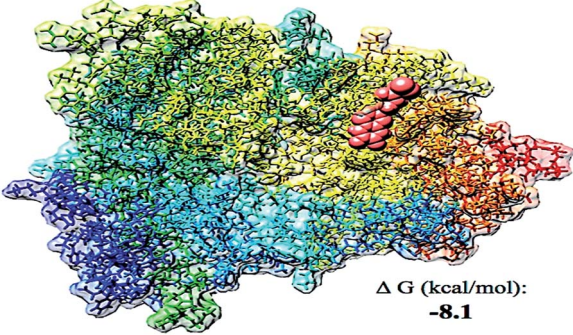
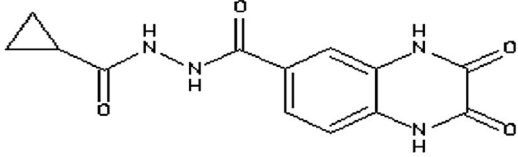
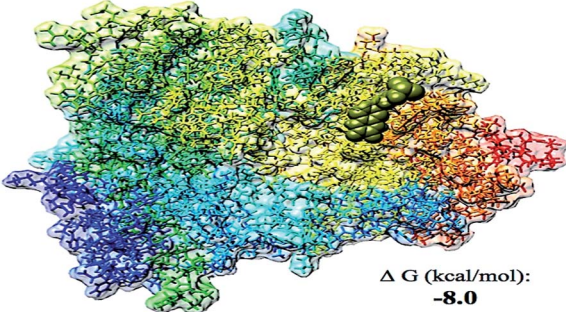
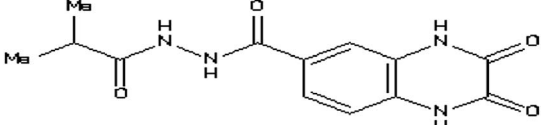
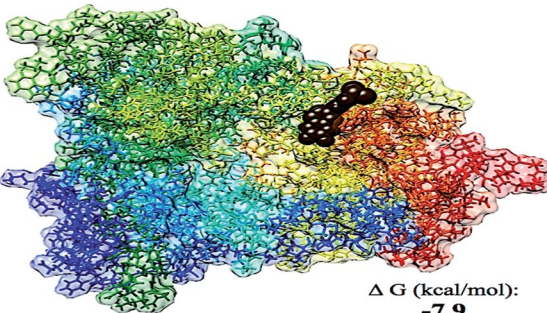
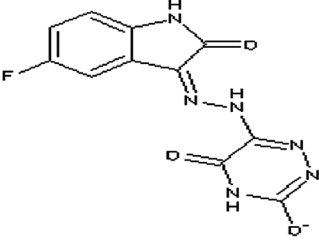
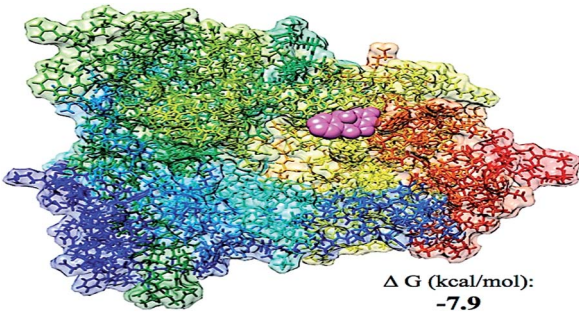
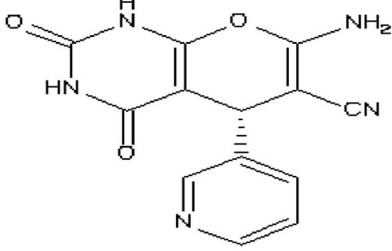
Table 2 Representation of top ten compounds docking to NS5 RdRp^a

Table 2 (Contd.)

 <p>ΔG (kcal/mol): -8.1</p>	 <p>ZINC13633807</p>
 <p>ΔG (kcal/mol): -8.1</p>	 <p>ZINC40563785</p>
 <p>ΔG (kcal/mol): -8.0</p>	 <p>ZINC40563886</p>
 <p>ΔG (kcal/mol): -7.9</p>	 <p>ZINC13121997</p>
 <p>ΔG (kcal/mol): -7.9</p>	 <p>ZINC00043707</p>

^a ΔG : binding affinity.

Table 3 Physical representation of top ten compounds displaying molecular weight, $x \log P$, H-bond donors/acceptors and rotatable bonds

ZINC ID	Molecular weight	$x \log P$	Rotatable bonds	H-Bond donors	H-Bond acceptors
ZINC35325271	291.27	−0.70	0	2	7
ZINC14987423	277.243	0.95	0	3	7
ZINC35325268	291.27	2.67	0	2	7
ZINC00351019	287.231	−5.47	0	5	8
ZINC13633807	299.246	−0.86	0	3	9
ZINC40563785	288.263	0.83	3	4	8
ZINC40563886	290.279	−0.89	3	4	8
ZINC13121997	289.206	−0.40	2	3	9
ZINC00043707	283.247	−1.01	1	4	8

model, where adult female mice were subcutaneously inoculated with similar doses of ZIKV from natural infection of a mosquito bite.^{67,68} This model is a critical cornerstone in accelerated testing of new ZIKV inhibitors.

Larroca *et al.* (2016) have made a significant contribution to the protection against the ZIKV virus by creating the first full-length prM-envelope DNA vaccine. The vaccine is currently undergoing clinical trials after the success in an *in vivo* study using infected mice. This vaccine may be the potential 'holo-grail' in ZIKV prevention.⁶⁹

5.1 A homology model for ZIKV NS5

In order for CADD of ZIKV to occur, a 3D crystal structure of a target protein is needed. Fig. 4 shows the first account of a homology model for the ZIKV NS5 protein, which was created and validated as described in our previous publications (Fig. 4).^{60,70} The PDB coordinates of the homology model are provided as.

The protein comprises of three domains, the N-terminal, methyltransferase domain (residues 1–262) (green), the inter-domain region (residues 263–272) (blue) and the C-terminal, RdRp domain (residues 273–903) (yellow).

5.2 Active site identification

Active site residues need to be identified for the docking of potential inhibitors to the active site pocket. The active site residues were determined using Chimera Multi-align Viewer and validated using the Site-Hound web program.⁷¹ Fig. 5 highlights the best active sites and active site residue numbers of the NS5 protein (Fig. 5).

This homology model will be implemented in the design of small molecules that may act as inhibitors of the NS5 protein, thus inhibiting the translation of viral RNA. Experimental drug therapy on other flaviviruses^{41,72,73} may be used as a guide toward the identification of new specified small molecules that inhibit ZIKV replication.

5.3 Possible small molecule inhibitors of NS5 RdRp

Of the ZIKV target proteins, NS5 RdRp is one of the most favorable for drug discovery due to its role in viral replication.¹ A study by Eyer *et al.* (2016) looked at an *in vitro* study of nucleoside inhibitors against ZIKV and found one particular molecule, 2-C-methyladenosine, to show promising inhibition of RdRp.⁷⁴ The purine and hydroxymethyl structural features of 2-C-

Table 4 Proposed computational software used in ZIKV drug design

Computational method	Software available	Software to be utilized in study
Homology modeling	Sequence alignment-Insight, Prime, Profit, LOOK, ICM, Sybyl, CLUSTALW Model construction-DS Modeller, Prime, LOOK, ICM, Sybyl, MODELLER, MOE, SWISS-MODEL, RaptorX, LOMETS, Phyre, I-Tasser	Sequence alignment-CLUSTALW ⁷⁷ Model construction-MODELLER ⁷⁸
Active binding site determination	CASTp, POOL, PASS, Pocket-Finder, 3DLigandSite, LIGSITE, metaPocket, FINDSITE, Site-hound	metaPocket ⁷⁹ POOL ⁸⁰ Site-hound ⁷¹
Molecular graphic systems	Avogadro, Chemlab, Athena, Maestro, Jmol, PyMOL, UCSF chimera, VMD, Vmol, Webmol, Zeus	UCSF chimera ¹²
Virtual screening databases	PubChem, MMSINC, ZINC, ZincPharmer, 4SC discovery, therapeutic target database, drug Bank, ChemSpider, ChEMBL	ZINC ⁸¹ ZincPharmer ⁸²
Docking software	PyRx, Autodock Vina, Dock Blaster, Vis3d, Schrodinger, GOLD, Libdock, FlexX, Glide, Fred, ICM	Autodock Vina ⁸³
Molecular dynamic simulation software	Gromacs, Amber, CHARM, Gromos, ADF, Desmond, NWChem	Amber ⁸⁴

methyladenosine were screened through ZINC database, criteria was imposed to ensure the inclusion of the maximum number of compounds, such that compounds had to have an $x \log P$ between -4 and 5 , a net charge 0 , rotatable bonds between 0 and 8 , a polar surface area of between 0 and 150 , have hydrogen bond donors/acceptors between 0 and 10 , and polar desolvation between 0 and 1 kcal mol^{-1} whereas compounds must have an apolar desolvation between -100 and 40 kcal mol^{-1} . Thereafter, the 4113 hits were downloaded and docked together with 2-C -methyladenosine (Fig. 6) at the RdRp active site and the nine best docked poses comprising the highest binding affinities were reported in Table 2. Table 3 shows the physical representation of the compounds. These compounds may be a basis for further validation and experimental verification.

Further information, including extensive procedures, can be found in our previous publications on structure-based enzymatic drug design.^{58,75} We believe that the robust computational tools implemented in the route map will provide a fundamental platform in the development of inhibitors against multiple ZIKV target molecules.

6. Proposed computational software that can be used in ZIKV drug design and discovery

The software available for techniques used in drug design have simplified the development of inhibitors allowing for specific binding to a target molecule, thus, decreasing its biological adverse effects.⁷⁶ There are various types of software available in structure-based drug design, allowing for faster and more comprehensive research into ZIKV inhibitors (Table 4).

7. Conclusion

The future of the ZIKV pandemic is uncertain and thus new, accelerated techniques are necessary to assist the medical and scientific community in the identification and validation of inhibitors to this global threat. The chemoinformatics discussed in this paper will not only in the identification and design of potential ZIKV inhibitors but also in parallel, but may assist in the early analysis of potential biological mutations that may occur due to the rapid international transmission of this *flavivirus*.

Conflicts of interest

Authors declare no potential conflicts of interest.

Acknowledgements

The authors acknowledge the College of Health Sciences, UKZN, and the National Research Foundation for their financial support.

References

- 1 O. Faye, C. C. M. Freire, A. Iamarino, O. Faye, J. V. C. de Oliveira, M. Diallo, *et al.*, Molecular Evolution of Zika Virus during Its Emergence in the 20th Century, *PLoS Neglected Trop. Dis.*, 2014, **8**(1), 1–10.
- 2 S. Bhakat, W. Karubiu, V. Jayaprakash and M. E. S. Soliman, A perspective on targeting non-structural proteins to combat neglected tropical diseases: Dengue, West Nile and Chikungunya viruses, *Eur. J. Med. Chem.*, 2014, **2014**(87), 677–702.
- 3 L. H. Chen and D. H. Hamer, Zika Virus: Rapid Spread in the Western Hemisphere, *Ann. Intern. Med.*, 2016, 1–3.
- 4 G. S. Campos, A. C. Bandeira and S. I. Sardi, Zika virus outbreak, Bahia, Brazil, *Emerging Infect. Dis.*, 2015, **21**(10), 1885–1886.
- 5 M. Mahfuz, A. Khan, M. H. Al, M. Hasan, A. Parvin, N. Rahman, *et al.*, Indian Journal of Pharmaceutical and Biological Research (IJPBR) *In Silico* Modeling and Immunoinformatics Probing Disclose the Epitope Based Peptide Vaccine Against Zika Virus Envelope Glycoprotein, *Indian J. Pharm. Biol. Res.*, 2014, **2**(4), 44–57.
- 6 M. Roa, Zika virus outbreak: Reproductive health and rights in Latin America, *Lancet*, 2016, **387**(10021), 843.
- 7 WHO, Zika virus, Microcephaly and Guillain-Barré, *Syndrome*, 2016, 1–12.
- 8 D. Musso, C. Roche, E. Robin, T. Nhan, A. Teissier and V. M. Cao-Lormeau, Potential sexual transmission of Zika virus, *Emerging Infect. Dis.*, 2015, **21**(2), 359–361.
- 9 S. Ekins, D. Mitchen, M. Coffee, T. Stratton, J. Freundlich, L. Freitas-Junior, *et al.*, Open drug discovery for the Zika virus, *F1000Research*, 2016, **5**, 150.
- 10 R. K. Singh, K. Dhama, Y. S. Malik, M. A. Ramakrishnan, K. Karthik, R. Tiwari, *et al.*, Zika Virus – emergence, evolution, pathology, diagnosis and control: current global scenario and future perspectives – a comprehensive review, *Vet. Q.*, 2016, **217**, 1–43.
- 11 D. Lissauer, E. Smit and M. D. Kilby, Zika virus and pregnancy, *Int. J. Gynecol. Obstet.*, 2016, 1258–1263.
- 12 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, *et al.*, UCSF Chimera – a visualization system for exploratory research and analysis, *J. Comput. Chem.*, 2004, **25**(13), 1605–1612.
- 13 A. Panchaud, M. Stojanov, A. Ammerdorffer and M. Vouga, Emerging Role of Zika Virus in Adverse Fetal and Neonatal Outcomes, *Clin. Microbiol. Rev.*, 2016, **29**(3), 659–694.
- 14 E. C. Chibueze, V. Tirado and O. Olukunmi, Zika virus infection in pregnancy: a systematic review of disease course and complications, *Bull. W. H. O.*, 2016, 1–35.
- 15 J. Mlakar, M. Korva, N. Tul, M. Popović, M. Poljšak-Prijatelj, J. Mraz, *et al.*, Zika Virus Associated with Microcephaly, *N. Engl. J. Med.*, 2016, **374**(10), 951–958.
- 16 J. M. Turmel, P. Amgueguen, B. Hubert, Y. M. Vandamme, M. Maquart, H. L. Guillou-Guillemette and I. L.-G. Art, Late sexual transmission of Zika virus related to, *Lancet*, 2016, **6736**(16), 30775.
- 17 A. S. Oliveira Melo, G. Malinger, R. Ximenes, P. O. Szejnfeld, S. Alves Sampaio and A. M. Bispo De Filippis, Zika virus intrauterine infection causes fetal brain abnormality and microcephaly: tip of the iceberg?, *Ultrasound Obstet. Gynecol.*, 2016, **47**(1), 6–7.

- 18 A. R. Plourde and E. M. Bloch, A Literature Review of Zika Virus, *Emerging Infect. Dis.*, 2016, **22**(7), 1–15.
- 19 G. Calvet, R. S. Aguiar, S. Melo, S. a Sampaio, I. de Filippis, A. Fabri, *et al.*, Case report of detection of Zika virus genome in amniotic fluid of affected fetuses: association with microcephaly outbreak in Brazil, *Lancet Infect. Dis.*, 2016, **3099**(16), 1–8.
- 20 L. R. Petersen, D. J. Jamieson, A. M. Powers and M. A. Honein, Zika Virus, *N. Engl. J. Med.*, 2016, **374**(16), 1552–1563.
- 21 J. J. Waggoner and B. A. Pinsky, Zika Virus: Diagnostics for an Emerging Pandemic Threat, *J. Clin. Microbiol.*, 2016, **54**(4), 860–867.
- 22 A. C. Gourinat, O. Connor, E. Calvez, C. Goarant and M. Dupont-Rouzeyrol, Detection of Zika virus in urine, *Emerging Infect. Dis.*, 2015, **21**(1), 84–86.
- 23 R. N. Charrel, I. Leparce-Goffart, S. Pas, X. de Lamballerie, M. Koopmans and C. Reuskens, State of knowledge on Zika virus for an adequate laboratory response, *Bull. W. H. O.*, 2016, 1–29.
- 24 A. D. Haddow, A. J. Schuh, C. Y. Yasuda, M. R. Kasper, V. Heang, R. Huy, *et al.*, Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage, *PLoS Neglected Trop. Dis.*, 2012, **6**(2), 1–7.
- 25 E. B. Hayes, Zika virus outside Africa, *Emerging Infect. Dis.*, 2009, **15**(9), 1347–1350.
- 26 G. W. a Dick, S. F. Kitchen and A. J. Haddow, Zika virus. I. Isolations and serological specificity, *Trans. R. Soc. Trop. Med. Hyg.*, 1952, **46**(5), 509–520.
- 27 M. R. Duffy, T.-H. Chen, W. T. Hancock, A. M. Powers, J. L. Kool, R. S. Lanciotti, *et al.*, Zika virus outbreak on Yap Island, Federated States of Micronesia, *N. Engl. J. Med.*, 2009, **360**(24), 2536–2543.
- 28 R. W. Malone, J. Homan, M. V. Callahan, J. Glasspool-Malone, L. Damodaran, A. D. B. Schneider, *et al.*, Zika Virus: Medical Countermeasure Development Challenges, *PLoS Neglected Trop. Dis.*, 2016, **10**(3), 1–26.
- 29 M. Basarab, C. Bowman, E. J. Aarons and I. Cropley, Zika virus, *BMJ*, 2016, **1049**, i1049.
- 30 J.-C. Saiz, Á. Vázquez-Calvo, A. B. Blázquez, T. Merino-Ramos, E. Escribano-Romero and M. A. Martín-Acebes, Zika Virus: the Latest Newcomer, *Front. Microbiol.*, 2016, **7**, 496.
- 31 S. Ekins, J. Liebler, B. J. Neves, W. G. Lewis, M. Coffee, R. Bienstock, *et al.*, Illustrating and homology modeling the proteins of the Zika virus, *F1000Research*, 2016, **5**, 275.
- 32 J. P. Hughes, S. S. Rees, S. B. Kalindjian and K. L. Philpott, Principles of early drug discovery, *Br. J. Pharmacol.*, 2011, **162**(6), 1239–1249.
- 33 C. G. Noble, Y. L. Chen, H. Dong, F. Gu, S. P. Lim, W. Schul, *et al.*, Strategies for development of dengue virus inhibitors, *Antiviral Res.*, 2010, **85**(3), 450–462.
- 34 R. Perera, M. Khaliq and R. J. Kuhn, Closing the door on flaviviruses: entry as a target for antiviral drug design, *Antiviral Res.*, 2008, **80**(1), 11–22.
- 35 K. J. Chappell, M. J. Stoermer, D. P. Fairlie and P. R. Young, Insights to substrate binding and processing by West Nile Virus NS3 protease through combined modeling, protease mutagenesis, and kinetic studies, *J. Biol. Chem.*, 2006, **281**(50), 38448–38458.
- 36 S. Agnihotri, R. Narula, K. Joshi, S. Rana and M. Singh, In silico modeling of ligand molecule for non structural 3 (NS3) protein target of flaviviruses, *Bioinformation*, 2012, **8**(3), 123–127.
- 37 H. Tian, X. Ji, X. Yang, W. Xie, K. Yang, C. Chen, *et al.*, The crystal structure of Zika virus helicase: basis for antiviral drug design, *Protein Cell*, 2016, **7**(6), 450–454.
- 38 H. Song, J. Qi, J. Haywood, Y. Shi and G. F. Gao, Zika virus NS1 structure reveals diversity of electrostatic surfaces among flaviviruses, *Nat. Struct. Mol. Biol.*, 2016, 1–4.
- 39 S. Youn, R. L. Ambrose, J. M. Mackenzie and M. S. Diamond, Non-structural protein-1 is required for West Nile virus replication complex formation and viral RNA synthesis, *Virology*, 2013, **10**, 339.
- 40 P. R. Beatty, H. Puerta-Guardo, S. S. Killingbeck, D. R. Glasner, K. Hopkins and E. Harris, Dengue virus NS1 triggers endothelial permeability and vascular leak that is prevented by NS1 vaccination, *Sci. Transl. Med.*, 2015, **7**(304), 304ra141.
- 41 H. J. Stahla-Beek, D. G. April, B. J. Saeedi, a. M. Hannah, S. M. Keenan and B. J. Geiss, Identification of a Novel Antiviral Inhibitor of the *Flavivirus* Guanylyltransferase Enzyme, *J. Virol.*, 2012, **86**(16), 8730–8739.
- 42 D. Sirohi, Z. Chen, L. Sun, T. Klose, T. C. Pierson, M. G. Rossmann, *et al.*, The 3.8 Å resolution cryo-EM structure of Zika virus, *Science*, 2016, **352**(6284), 467–470.
- 43 M. Courageot, M. Frenkiel, S. D. Dos, V. Deubel, P. Desprès, C. Duarte, *et al.*, α -Glucosidase Inhibitors Reduce Dengue Virus Production by Affecting the Initial Steps of Virion Morphogenesis in the Endoplasmic Reticulum – Glucosidase Inhibitors Reduce Dengue Virus Production by Affecting the Initial Steps of Virion Morphogenesis in, *J. Virol.*, 2000, **74**(1), 564–572.
- 44 R. Hamel, O. Dejarnac, S. Wichit, P. Ekchariyawat, A. Neyret, N. Luplertlop, *et al.*, Biology of Zika Virus Infection in Human Skin Cells, *J. Virol.*, 2015, **89**(17), 8880–8896.
- 45 W. Noppakunmongkolchai, T. Poyomtip, T. Jittawuttipoka, N. Luplertlop, A. Sakuntabhai, S. Chimnaronk, *et al.*, Inhibition of protein kinase C promotes dengue virus replication, *Virology*, 2016, **13**(1), 35.
- 46 A. J. Hirsch, In Viral Replication, *Future Med. Chem.*, 2010, **5**(2), 303–311.
- 47 R. Zhang, J. J. Miner, M. J. Gorman, K. Rausch, H. Ramage, J. P. White, *et al.*, A CRISPR screen defines a signal peptide processing pathway required by flaviviruses [Internet], *Nature*, 2016, 164–168.
- 48 M. Mahfuz, K. Ali, H. A. Mahmud, H. Mahmudul, A. Parvin, R. Nazibur and S. M. Badier, Design and Prediction of Potential RNAi (siRNA) Molecules for 3' UTR PTGS of different strains of Zika Virus: A Computational Approach, *Nat. Sci.*, 2015, **1**(2), 37–50.
- 49 V. A. Kostyuchenko, E. X. Y. Lim, S. Zhang, G. Fibriansah, T.-S. Ng, J. S. G. Ooi, *et al.*, Structure of the thermally stable Zika virus, *Nature*, 2016, **533**(7603), 425–428.

- 50 L. Dai, J. Song, X. Lu, Y.-Q. Deng, A. M. Musyoki, H. Cheng, *et al.*, Structures of the Zika Virus Envelope Protein and Its Complex with a *Flavivirus* Broadly Protective Antibody, *Cell Host Microbe*, 2016, **19**(5), 696–704.
- 51 G. Barba-Spaeth, W. Dejnirattisai, A. Rouvinski, M.-C. Vaney, I. Medits, A. Sharma, *et al.*, Structural basis of potent Zika-dengue virus antibody cross-neutralization, *Nature*, 2016, 1–23.
- 52 J. K. Weltman, Medical Microbiology & Diagnosis An Immuno-Bioinformatic Analysis of Zika virus (ZIKV) Envelope E Protein, *J. Med. Microbiol. Diagn.*, 2016, **5**(2), 1–2.
- 53 S. Mandal, M. Moudgil and S. K. Mandal, Rational drug design, *Eur. J. Pharmacol.*, 2009, **625**(1–3), 90–100.
- 54 C. M. Song, S. J. Lim and J. C. Tong, Recent advances in computer-aided drug design, *Briefings Bioinf.*, 2009, **10**(5), 579–591.
- 55 P. Bamborough and F. E. Cohen, Modeling protein-ligand complexes, *Curr. Opin. Struct. Biol.*, 1996, **6**(2), 236–241.
- 56 A. C. Anderson, The Process of Structure-Based Drug Design, *Chem. Biol.*, 2003, **10**(9), 787–797.
- 57 H. J. Huang, H. W. Yu, C. Y. Chen, C. H. Hsu, H. Y. Chen, K. J. Lee, *et al.*, Current developments of computer-aided drug design, *J. Taiwan Inst. Chem. Eng.*, 2010, **41**(6), 623–635.
- 58 B. Honarparvar, T. Govender, G. E. M. Maguire, M. E. S. Soliman and H. G. Kruger, Integrated Approach to Structure-Based Enzymatic Drug Design: Molecular Modeling, Spectroscopy, and Experimental Bioactivity, *Chem. Rev.*, 2014, **114**, 493–537.
- 59 R. W. Hooft, C. Sander and G. Vriend, Objectively judging the quality of a protein structure from a Ramachandran plot, *Computer Applications in the Biosciences*, 1997, **13**(4), 425–430.
- 60 S. Chetty and M. E. S. Soliman, Possible allosteric binding site on Gyrase B, a key target for novel anti-TB drugs: homology modelling and binding site identification using molecular dynamics simulation and binding free energy calculations, *Med. Chem. Res.*, 2015, **2015**(24), 2055–2074.
- 61 S. M. Saberi Fathi and J. a. Tuszynski, A simple method for finding a protein's ligand-binding pockets, *BMC Struct. Biol.*, 2014, **14**(18), 1–9.
- 62 E. Lionta, G. Spyrou, D. K. Vassilatis and Z. Cournia, Structure-based virtual screening for drug discovery: principles, applications and recent advances, *Curr. Top. Med. Chem.*, 2014, **14**(16), 1923–1938.
- 63 M. Ramesh, S. B. Vepuri, F. Oosthuizen and M. E. Soliman, Adenosine Monophosphate-Activated Protein Kinase (AMPK) as a Diverse Therapeutic Target: A Computational Perspective, *Appl. Biochem. Biotechnol.*, 2015, 1–21.
- 64 S. S. Reddy, P. P. Pati, P. P. Kumar, H. Pradeep and N. N. Sastry, Virtual screening in drug discovery – a computational perspective, *Curr. Protein Pept. Sci.*, 2007, **8**(4), 329–351.
- 65 G. Sliwoski, S. Kothiwale, J. Meiler and E. W. Lowe, Computational methods in drug discovery, *Pharmacol. Rev.*, 2014, **66**(1), 334–395.
- 66 H. M. Kumalo and M. E. Soliman, Per-Residue Energy Footprints-Based Pharmacophore Modeling as an Enhanced *In Silico* Approach in Drug Discovery: A Case Study on the Identification of Novel β -Secretase1 (BACE1) Inhibitors as Anti-Alzheimer Agents, *Cell. Mol. Bioeng.*, 2015, **9**(1), 175–189.
- 67 R. Delvecchio, L. M. Higa, P. Pezzuto, A. L. Valadao, P. P. Garcez, F. L. Monteiro, *et al.*, Chloroquine inhibits Zika virus infection in different cellular models, *bioRxiv, Biochem.*, 2016, 051268.
- 68 S. D. Dowall, V. A. Graham, E. Rayner, B. Atkinson, G. Hall, R. J. Watson, *et al.*, A Susceptible Mouse Model for Zika Virus Infection, *PLoS Neglected Trop. Dis.*, 2016, **10**(5), e0004658.
- 69 R. A. Larocca, P. Abbink, J. P. S. Peron, P. M. A. Zanotto de, M. J. Iampietro, A. Badamchi-Zadeh, *et al.*, Vaccine protection against Zika virus from Brazil, *Nature*, 2016, 1–8.
- 70 Y. Maharaj and M. E. S. Soliman, Identification of novel gyrase b inhibitors as potential Anti-TB drugs: Homology modelling, hybrid virtual screening and molecular dynamics simulations, *Chem. Biol. Drug Des.*, 2013, **82**(2), 205–215.
- 71 M. Hernandez, D. Ghersi and R. Sanchez, SITEHOUND-web: A server for ligand binding site identification in protein structures, *Nucleic Acids Res.*, 2009, **37**(2), 413–416.
- 72 U. S. F. Tambunan, H. Zahroh, B. B. Utomo and A. A. Parikesit, Screening of commercial cyclic peptide as inhibitor NS5 methyltransferase of Dengue virus through Molecular Docking and Molecular Dynamics Simulation, *Bioinformation*, 2014, **10**(1), 23–27.
- 73 P. Niyomrattanakit, Y. L. Chen, H. Dong, Z. Yin, M. Qing, J. F. Glickman, *et al.*, Inhibition of dengue virus polymerase by blocking of the RNA tunnel, *J. Virol.*, 2010, **84**(11), 5678–5686.
- 74 L. Eyer, R. Nencka, I. Huvarova, M. Palus, M. Alves, E. A. Gould, *et al.*, Nucleoside inhibitors of Zika virus, *J. Infect. Dis.*, 2016, **27**(May), 1–13.
- 75 S. Moonsamy, S. Bhakat and M. E. S. Soliman, Dynamic features of apo and bound HIV-Nef protein reveal the anti-HIV dimerization inhibition mechanism, *J. Recept. Signal Transduction*, 2014, **9893**, 1–11.
- 76 C. Liao, M. Sitzmann, A. Pugliese and M. C. Nicklaus, Software and resources for computational medicinal chemistry, *Future Med. Chem.*, 2011, **3**(8), 1057–1085.
- 77 J. D. Thompson, T. J. Gibson and D. G. Higgins, Multiple sequence alignment using ClustalW and ClustalX, *Current Protocols in Bioinformatics*, 2002, **2**(2.3), 1–22.
- 78 N. Eswar, B. Webb, M. a Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, *et al.*, Comparative protein structure modeling using Modeller [Internet], *Current Protocols in Bioinformatics*, 2006, 1–47.
- 79 S. Somarowthu, H. Yang, D. G. C. Hildebrand and M. J. Ondrechen, High-performance prediction of functional residues in proteins with machine learning and computed input features, *Biopolymers*, 2011, **95**(6), 390–400.
- 80 B. Huang, MetaPocket: a meta approach to improve protein ligand binding site prediction, *OMICS*, 2009, **13**(4), 325–330.

- 81 J. J. Irwin and B. K. Shoichet, ZINC – a free database of commercially available compounds for virtual screening, *J. Chem. Inf. Model.*, 2005, **45**(1), 177–182.
- 82 D. R. Koes and C. J. Camacho, ZINCPharmer: Pharmacophore search of the ZINC database, *Nucleic Acids Res.*, 2012, **40**(W1), 409–414.
- 83 O. Trott and A. J. Olson, AutoDock Vina, *J. Comput. Chem.*, 2010, **31**, 445–461.
- 84 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, *et al.*, The Amber biomolecular simulation programs, *J. Comput. Chem.*, 2005, 1668–1688.

Zika Virus Drug Targets: A Missing Link In Drug Design And Discovery – A Route Map To Fill The Gap

Pritika Ramharack^a and Mahmoud E. S. Soliman^{a*}

Supplementary Material

Figure S1: The multiple sequence alignment

```

gi|534286613|pdb|4K6M|A      -GRPGGRTLGEQWKEKLNAMSREEFFKYRREAIIEVDRTTEARRARRENNI
gi|126030613|pdb|2HFZ|A      -----
gi|985757036|ref|YP_009227205. -GGGTGETLGEKWKARLNQMSALEFYKYKSGITEVCREEARRALKDQVA
gi|985483961|pdb|5CCV|A      GTGSQGETLGEKWKKNLQLSRKEFDLYKKSGITEVDRTTEAKEGLKRGEI
gi|453055587|pdb|3VWS|A      -----

gi|534286613|pdb|4K6M|A      VGGHPVSRGSAKLRLVLEKGFVSPIGKVIDLGCGRGGWSYYAATLKKVQE
gi|126030613|pdb|2HFZ|A      -----
gi|985757036|ref|YP_009227205. TGGHAVSRGSAKIRWLEERGYPYGVVDLGCGRGGWSYYAATIRKVVQE
gi|985483961|pdb|5CCV|A      T-HHAVSRGSAKLQWFVERNMVPEGRVIDLGCGRGGWSYYCAGLKKVTE
gi|453055587|pdb|3VWS|A      -----

gi|534286613|pdb|4K6M|A      VRGYTKGGAGHEEPMLMQSYGWNVLVSLKSGVDVFFYKPEPSDTLFCDIGE
gi|126030613|pdb|2HFZ|A      -----
gi|985757036|ref|YP_009227205. VRGYTKGGPGHEEPMLVQSYGWNIVRLKSGVDVFHMAAEPDCTLCDIGE
gi|985483961|pdb|5CCV|A      VRGYTKGGPGHEEPVPMSTYGVNIVKLMMSGKDVFFYLPFEKCDTLCDIGE
gi|453055587|pdb|3VWS|A      -----

gi|534286613|pdb|4K6M|A      SSPSPVEVEEQRTLRLVLEMTSDWLHRGPREFCIKVLCPYMPKVKIEKMEVLQ
gi|126030613|pdb|2HFZ|A      -----
gi|985757036|ref|YP_009227205. SSSSPVEVEETRTLRLVLSMVGDWLEKRPAGFCIKVLCPYTSTMMETMERLQ
gi|985483961|pdb|5CCV|A      SSPSPVTEESRTIRVLKMVEPWLKN--NQFCIKVLNPMPTVIEHLERLQ
gi|453055587|pdb|3VWS|A      -----

gi|534286613|pdb|4K6M|A      RRFGGGLVRLPLSRNSNHEMYWVWSGAAGNVVHAVNMTSQVLLGRMDRTVW
gi|126030613|pdb|2HFZ|A      -----
gi|985757036|ref|YP_009227205. RRHGGGLVVRVPLCRNSTHEMYWVSGAKSNIIKSVSTTSQVLLGRMD-GPR
gi|985483961|pdb|5CCV|A      RKHGGMLVRNPLSRNSTHEMYWISNGTGNIVSSVNMVSRLLLNRFMTMTHR
gi|453055587|pdb|3VWS|A      -----

gi|534286613|pdb|4K6M|A      RGPKYEEDVNLGSGTRAVGKGEVHSNQEKIKKRIQKLKEEFATTWHKDPE
gi|126030613|pdb|2HFZ|A      -----
gi|985757036|ref|YP_009227205. -----HHHHHHKSDTSKIKNRIERLRREYSSTWHHDEN
gi|985483961|pdb|5CCV|A      RPVKYEEEDVNLGSGTRAVASCAEAPNMKIIIGRIERIRNEHAETWFLDEN
gi|453055587|pdb|3VWS|A      RP-TIEKDVDLGAAGRHRVNAEPETPNMDVIGERIKRIKEEHSSTWHYDDE
-----GSHMLDNMDVIGERIKRIKEEHNSTWHYDDE
          : * . * : : : : . * . * . * :

gi|534286613|pdb|4K6M|A      HPYRTWYTHGSEYVKATGSASSLVNGVVKLSKPWDAIANVTTMAMTDTT
gi|126030613|pdb|2HFZ|A      HPYRTWYHGSYEVKPTGSASSLVNGVVRLLSKPWDITITNVTMTAMTDTT
gi|985757036|ref|YP_009227205. HPYRTWAYHGSYEAPTQGSASSLVNGVVRLLSKPWDVVTGVTGIAMTDTT
gi|985483961|pdb|5CCV|A      NPYKTWAYHGSYEVKATGSASSMINGVVKLLTKPVDVVPMTQAMTDTT
gi|453055587|pdb|3VWS|A      NPYKTWAYHGSYEVKATGSASSMINGVVKLLTKPVDVVPMTQAMTDTT
: * : * * * * * . . * * * : : * * : : * * : . * * : * * * *

gi|534286613|pdb|4K6M|A      PFGQQRVFKEKVDTKAPEPPAGAKEVLNETTNWLWAYLSREKRPRCLCTKE
gi|126030613|pdb|2HFZ|A      PFGQQRVFKEKVDTKAPEPPPEGVYVNLNETTNWLWAFLAEREKPRMCSRE
gi|985757036|ref|YP_009227205. PYGQQRVFKEKVDTRVPDPQEGTRQVMNIVSSWLWKLGRKRPRVCTKE
gi|985483961|pdb|5CCV|A      PFGQQRVFKEKVDTRTPRPMGTRKVMETAEWLWRTLGRNKRPRCLCTRE
gi|453055587|pdb|3VWS|A      PFGQQRVFKEKVDTRTPRPLPGTRKVMETAEWLWRTLGRNKRPRCLCTRE
* : * * * * * * * * : . * * * . : : : . : * * * . : . : * * * : : *

gi|534286613|pdb|4K6M|A      EFIRKVNNSNAALGAVFAEQNQWSTAREAVDDPRFWEMVDEERENHLRGEC
gi|126030613|pdb|2HFZ|A      EFIRKVNNSNAALGAMFEEQNQWRSAREAVEDPKFWEMVDEEREALHREGC
gi|985757036|ref|YP_009227205. EFINKVRSNAALGAI FEEKEWKTAVEAVNDPRFWALVDREREHHLRGEC
gi|985483961|pdb|5CCV|A      EFTKKVRTNAAMGAVFTEENQWDSARA AVEDEEFWKLVDRERELHKLKGK
gi|453055587|pdb|3VWS|A      EFTKKVRTNAAMGAVFTEENQWDSAKAAVEDEEFWKLVDRERELHKLKGK
** . * . : * * : * * : * * : : : * * * : * . * * : * * : * * : *

gi|534286613|pdb|4K6M|A      HTCIYNMGMKREKKPGFEGKAKGSRAIWFMWLGARYLEFEALGFLNEDHW
gi|126030613|pdb|2HFZ|A      HTCIYNMGMKREKKPGFEGKAKGSRAIWFMWLGARFLEFEALGFLNEDHW
gi|985757036|ref|YP_009227205. HSCVYNNMGMKREKKQGEFGKAKGSRAIWMWLGARFLEFEALGFLNEDHW
gi|985483961|pdb|5CCV|A      GSCVYNNMGMKREKKLGEFGKAKGSRAIWMWLGARYLEFEALGFLNEDHW
gi|453055587|pdb|3VWS|A      GSCVYNNMGMKREKKLGEFGKAKGSRAIWMWLGVRYLEFEALGFLNEDHW
: * : * * * * * * * * * * * * : * * * . * : * * * * * * * *

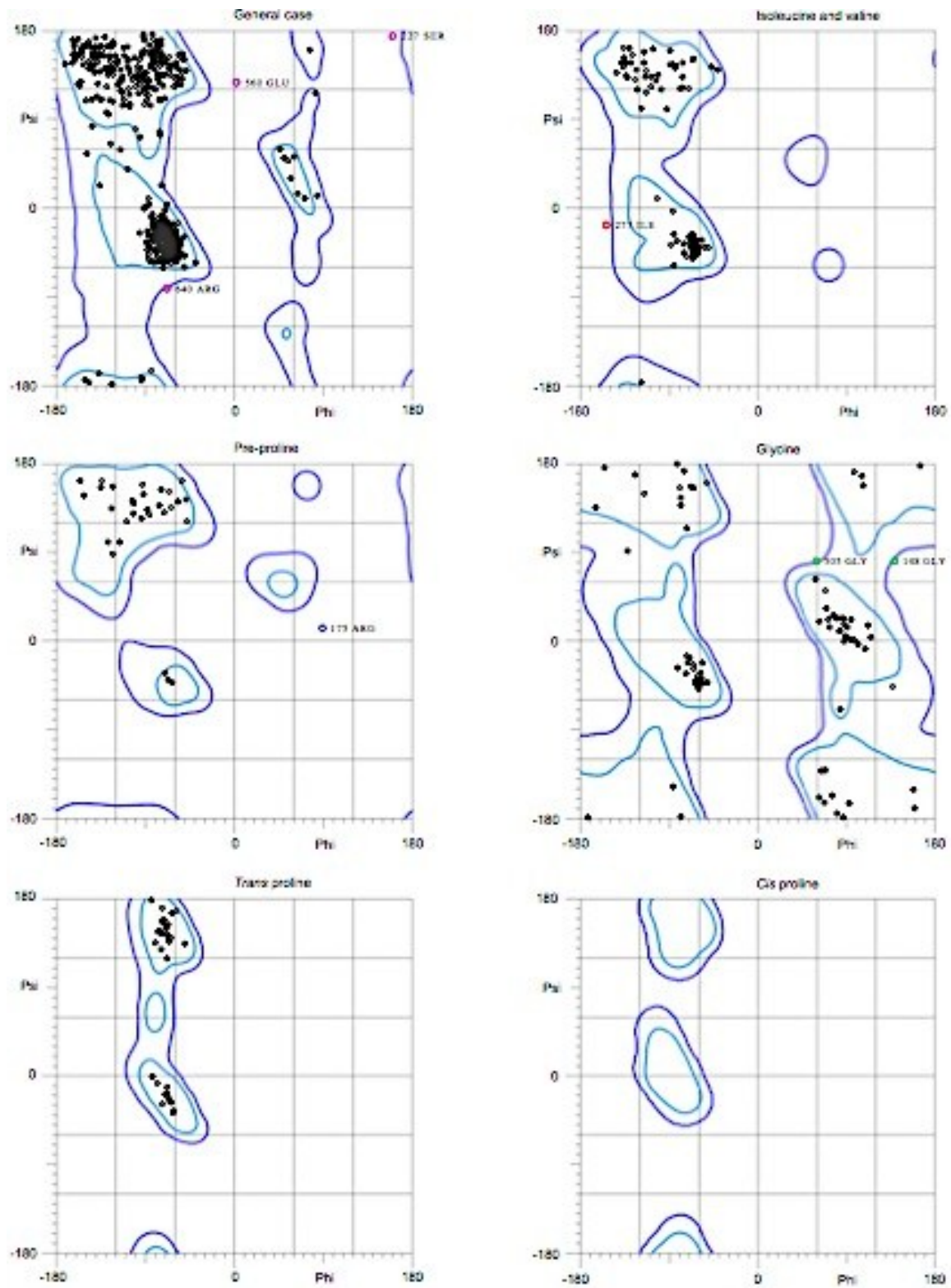
gi|534286613|pdb|4K6M|A      LSRENSGGGVEGSGVQKLGYYILRDIAGKQGGKMYADDTAGWDTRITRIDL
gi|126030613|pdb|2HFZ|A      LGRKNSGGGVEGLGLQKLGYYILREVGTTPGGRIYADDTAGWDTRITRADL
gi|985757036|ref|YP_009227205. MGRNSGGGVEGLGLQRLGYILEEMNRAPGGKMYADDTAGWDTRI SKFDL
gi|985483961|pdb|5CCV|A      FSRENSYSGVEGEGLHKLGYILRDISKIPGGAMYADDTAGWDTRITEDDL
gi|453055587|pdb|3VWS|A      FSRENSYSGVEGEGLHKLGYILRDISKIPGGAMYADDTAGWDTRITEDDL
: . * : * * . * * * * : : : * * * * * * * * : . * * : * *

gi|534286613|pdb|4K6M|A      ENEAKVLELLDGEHRMLARAI IELTYRHKVVVKVMPRAAEGKTVMDVISRE
gi|126030613|pdb|2HFZ|A      ENEAKVLELLDGEHRLARAI IELTYRHKVVVKVMPRAADGRTVMDVISRE
gi|985757036|ref|YP_009227205. ENEALITNQMEEGHRTLALAVIKYTYQNKVVVKVLRPAEGGKTVMDIISRK
gi|985483961|pdb|5CCV|A      HNEEKITQQMDPEHRQLANAI FKLTYQNKVVVKVQRPTPKG-TVMDIISRK
gi|453055587|pdb|3VWS|A      HNEEKI IQQMDPEHRQLANAI FKLTYQNKVVVKVQRPTPTG-TVMDIISRK
. * * : : : : * * * * : : : * * : * * * * * * : * * * * *

```


gi 534286613 pdb 4K6M A	DQRGSGQVVITYALNTFTNIAVQLVRLMEAEVIGPQHLEQLPRKNKIAVR
gi 126030613 pdb 2HFZ A	DQRGSGQVVITYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTGKGPKVR
gi 985757036 ref YP_009227205.	DQRGSGQVVITYALNTFTNLVQLIRNMEAEVLEMQDLWLLRKPE--KVT
gi 985483961 pdb 5CCV A	DQRGSGQVGTYGLNTFTNMEAQLIRQMEGEGVLSKADLENP-HPLEKKIT
gi 453055587 pdb 3VWS A	DQRGSGQVGTYGLNTFTNMEAQLVRQMEGEGVLTAKADLENP-HLLEKKIT
	***** **.******: .**:* **.* *: .: : *
gi 534286613 pdb 4K6M A	TWLFENGEERVTRMAISGDDCVVKPLDDRFATALHFLNAMS KVRKDIQEW
gi 126030613 pdb 2HFZ A	TWLFENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMS KVRKDIQEW
gi 985757036 ref YP_009227205.	RWLQSNQWDRLLKRMVSGDDCVVKPIDDRFAHALRFLNDMGKVRKDTQEW
gi 985483961 pdb 5CCV A	QWLETKGVERLKRMAISGDDCVVKPIDDRFANALLALNDMGKVRKDIQW
gi 453055587 pdb 3VWS A	QWLETKGVERLKRMAISGDDCVVKPIDDRFANALLALNDMGKVRKDIQW
	** :* :* :*.***:*****:***** :* ** *.***** :*
gi 534286613 pdb 4K6M A	KPSHGWHWDWQQVPFCSNHFEIVMKDGRSIVVPCRQDELIGRARISPGA
gi 126030613 pdb 2HFZ A	KPSTGWDYDQQVPFCSNHFEIIMKDGRITLVTPCRGQDELIVGRARISPGA
gi 985757036 ref YP_009227205.	KPSTGWSNWEVFPFCSHHFNKLYLKDGRSIVVPCRQDELIGRARISPGA
gi 985483961 pdb 5CCV A	QPSKGWHWDWQQVPFCSHHFHELIMKDGRKLVVPCRQDELIGRARISQGA
gi 453055587 pdb 3VWS A	QPSKGWHWDWQQVPFCSHHFHELIMKDGRKLVVPCRQDELIGRARISQGA
	:** ** :*:*****:*** : : :*****. :*.*** *****:*** **
gi 534286613 pdb 4K6M A	GWNVKTACIAKAYAQMWLLLYFHRRDLRLMANAICSAVPVDWVPTGRTS
gi 126030613 pdb 2HFZ A	GWNVRDTACIAKASYAQMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRTT
gi 985757036 ref YP_009227205.	GWSIRETACIAKASYAQMWQLLYFHRRDLRLMANAICSAVPVDWVPTGRTT
gi 985483961 pdb 5CCV A	GWSLRETACIAKAYAQMWALMYFHRRDLRLASNAICSAVPVHWVPTSR--
gi 453055587 pdb 3VWS A	GWSLRETACIAKAYAQMWALMYFHRRDLRLASNAICSAVPVHWVPTSR--
	.: :***. :*.***** :***** :*****.*****. *
gi 534286613 pdb 4K6M A	WSIHSGGEWMTTEDMLQVWNRVWIEENEWMMDKTPITSWTDVPYVGKRED
gi 126030613 pdb 2HFZ A	WSIHAGGEWMTTEDMLEVWNRVWIEENEWMEDKTPVEKWSDVYPYSGKRED
gi 985757036 ref YP_009227205.	WSIHGSGGEWMTTEDMLMVWNRVWIEENDHMDKTPVTKTWDIPYLGKRED
gi 985483961 pdb 5CCV A	----TTHQWMTTEDMLTVWNRVWIEDNPWMEDKTPVTTWEDVPYLGKRED
gi 453055587 pdb 3VWS A	WSIHAAHQWMTTEDMLTVWNRVWIEENPWMDKTPVTTWENVPYLGKRED
	:***** *****:*** * ****: .* : :*** *****
gi 534286613 pdb 4K6M A	IWCGSLIGTRSRATWAENIYAAINQVRAVIGKE-NYVDYMTSLRRYEDVL
gi 126030613 pdb 2HFZ A	IWCGSLIGTRARATWAENIQVAINQVRSIIGDE-KYVDYMSLKRKYEDTT
gi 985757036 ref YP_009227205.	LWCGSLIGHRPRTTWAENIKDTVNMVRRIGDEEKYMDYLSQVRYLGEE
gi 985483961 pdb 5CCV A	QWCGSLIGLTSRATWAQNIPTAIQVRSIIGNE-EFLDYMPMSMKRFRKEE
gi 453055587 pdb 3VWS A	QWCGSLIGLTSRATWAQNIPTAIQVRSIIGNE-EFLDYMPMSMKRFRKEE
	***** .*:*****:*** : : ** :***. * : :*****. : *
gi 534286613 pdb 4K6M A	IQEDRVIGSSSHHHHHH-
gi 126030613 pdb 2HFZ A	LVEDTVL-----
gi 985757036 ref YP_009227205.	GSTPGVL-----
gi 985483961 pdb 5CCV A	ESEGAIWAAALEHHHHHH
gi 453055587 pdb 3VWS A	ESEGAIW-----

Figure S2: Ramachandran plot for ZIKV NS5 protein

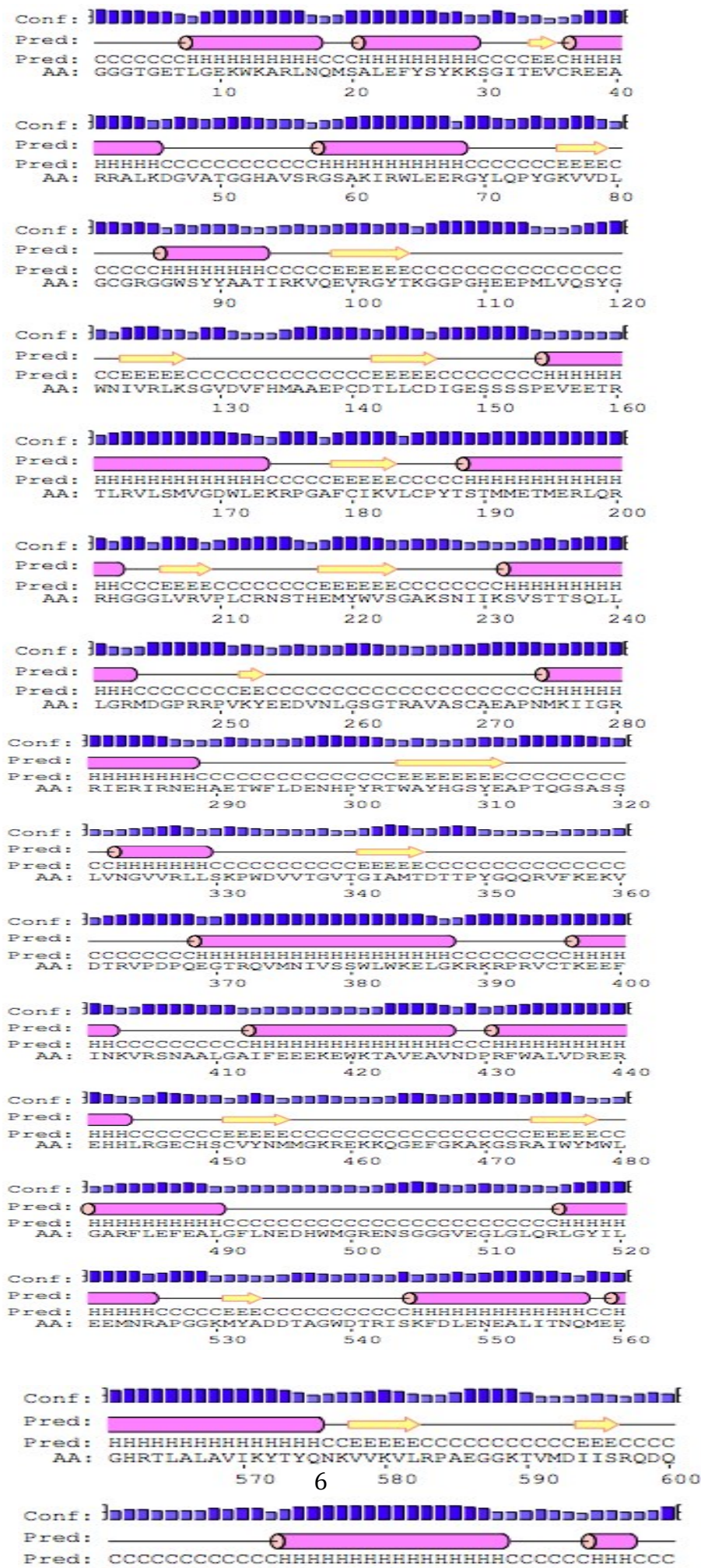


97.2% (876/901) of all residues were in favored (98%) regions.
99.2% (894/901) of all residues were in allowed (>99.8%) regions.

There were 7 outliers (phi, psi):

148 GLY (138.4, 81.3)
175 ARG (89.6, 13.4)
227 SER (161.0, 175.0)
277 ILE (-153.6, -17.7)
505 GLY (59.8, 81.0)
560 GLU (2.5, 127.7)
640 ARG (-68.3, -81.5)

Figure S3: Psipred predicted secondary structure





Zika virus NS5 protein potential inhibitors: an enhanced in silico approach in drug discovery

Pritika Ramharack & Mahmoud E.S. Soliman

To cite this article: Pritika Ramharack & Mahmoud E.S. Soliman (2017): Zika virus NS5 protein potential inhibitors: an enhanced in silico approach in drug discovery, Journal of Biomolecular Structure and Dynamics, DOI: [10.1080/07391102.2017.1313175](https://doi.org/10.1080/07391102.2017.1313175)

To link to this article: <http://dx.doi.org/10.1080/07391102.2017.1313175>



View supplementary material [↗](#)



Accepted author version posted online: 29 Mar 2017.
Published online: 17 Apr 2017.



Submit your article to this journal [↗](#)



Article views: 37



View related articles [↗](#)



View Crossmark data [↗](#)



Zika virus NS5 protein potential inhibitors: an enhanced *in silico* approach in drug discovery

Pritika Ramharack^a and Mahmoud E.S. Soliman^{a,b,c,d,*}

^aMolecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa; ^bPharmaceutical Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa; ^cFaculty of Pharmacy, Department of Pharmaceutical Organic Chemistry, Zagazig University, Zagazig, Egypt; ^dCollege of Pharmacy and Pharmaceutical Sciences, Florida Agricultural and Mechanical University, FAMU, Tallahassee, FL 32307, USA

Communicated by Ramaswamy H. Sarma

(Received 30 December 2016; accepted 22 March 2017)

The re-emerging Zika virus (ZIKV) is an arthropod-borne virus that has been described to have explosive potential as a worldwide pandemic. The initial transmission of the virus was through a mosquito vector, however, evolving modes of transmission has allowed the spread of the disease over continents. The virus has already been linked to irreversible chronic central nervous system conditions. The concerns of the scientific and clinical community are the consequences of Zika viral mutations, thus suggesting the urgent need for viral inhibitors. There have been large strides in vaccine development against the virus but there are still no FDA approved drugs available. Rapid rational drug design and discovery research is fundamental in the production of potent inhibitors against the virus that will not just mask the virus, but destroy it completely. *In silico* drug design allows for this prompt screening of potential leads, thus decreasing the consumption of precious time and resources. This study demonstrates an optimized and proven screening technique in the discovery of two potential small molecule inhibitors of ZIKV Methyltransferase and RNA dependent RNA polymerase. This *in silico* 'per-residue energy decomposition pharmacophore' virtual screening approach will be critical in aiding scientists in the discovery of not only effective inhibitors of Zika viral targets, but also a wide range of anti-viral agents.

Keywords: Zika virus per-residue decomposition based pharmacophore; virtual screening; NS5 protein potential inhibitors; binding free energy; molecular dynamic simulations

1. Introduction

Zika virus (ZIKV) is an arthropod-borne virus that has been described to have potential as a worldwide pandemic (Troncoso, 2016). The virus is a member of the *spodweni* serocomplex of the *flavivirus* genus and was first discovered in 1947 by its isolation from the Rhesus 766 monkey in Uganda (Faye et al., 2014; Haddow et al., 2012). Sporadic cases of the virus have been reported in countries such as Uganda, Tanzania, Egypt, Gabon, and in parts of Asia including India and Indonesia, with the most devastating epidemic occurring in Brazil in 2015 (Campos, Bandeira, & Sardi, 2015; Mahfuz et al., 2015). As of June 2016, 11 countries had reported central nervous system (CNS) malformations potentially linked to ZIKV. During 2015 and early 2016, eight countries had reported cases of Guillian-Barré syndrome (GBS), where laboratory testing confirmed ZIKV infection was found in a number of GBS cases (WHO, 2016).

Transmission of the virus was thought to be only via the *Aedes* mosquito vector but studies during 2016, have evidenced congenital, perinatal, and sexual transmission (Singh et al., 2016; Turmel, Hubert, Maquart,

Guillou-Guillemette, & Leparac-Goff, 2016). The virus triggers febrile like influenza conditions in the host, including swollen lymph nodes, skin rashes, and joint pains (Brito, 2016; Ekins et al., 2016; Shapshak, Sinnott, Somboonwit, & Kuhn, 2016). The concerns of the scientific community involve the dramatic increase in ZIKV-related CNS disorders including neonatal-microcephaly and GBS (Lissauer, Smit, & Kilby, 2016; Panchaud, Stojanov, Ammerdorffer, & Vouga, 2016; Roa, 2016). Complications associated with prenatal infection encompass fetal growth restriction, neurological and ocular abnormalities, intracranial calcification and in some cases perinatal death or stillbirth (Chibueze et al., 2017; Singh et al., 2016).

The virus is able to enter a host via receptor-mediated endocytosis, followed by fusion from within the endosomal cell compartment (Mahfuz et al., 2014). The enveloped virus comprises an 11 kilo base, single-stranded positive sense RNA genome which consists of 10,794 nucleotides encoding 3419 amino acids (Hayes, 2009). The open reading frame of the 5' and 3' untranslated region encodes a polyprotein cleaved into three

*Corresponding author. Email: soliman@ukzn.ac.za

structural proteins being the capsid, premembrane/membrane, and envelope. Seven non-structural proteins may also be found in this assembly, namely, NS1, NS2A, NS2B, NS3, NS4A, 2K, NS4B, and NS5 (largest viral protein) (Haddow et al., 2012).

Being the largest and most imperative protein in the genome replication and RNA capping of ZIKV, NS5 presents as a novel antiviral target (Tambunan, Zahroh, Utomo, & Parikesit, 2014). The protein consists of three domains: a Methyltransferase (MTase) domain at residues 1–262 of its N-terminal, an RNA dependent RNA polymerase (RdRp) at residues 273–903 of its C-terminal and an inter-domain region at residues 263–272 (Zou et al., 2014).

The MTase domain belongs to the family of S-Adenosyl Methionine (SAM)-dependent enzymes, containing a SAM-dependent MTase fold comprising of an $\alpha/\beta/\alpha$ structure (Zou et al., 2014). The MTase domain is one of the key targets in drug design as the enzyme performs nucleoside-2'O and N-7 methylation of the viral RNA cap which is essential in the replication of the virus (Egloff, Benarroch, Selisko, Romette, & Canard, 2002). Upon the completion of methylation, SAM is converted to S-Adenosyl Homocysteine and gets released from the MTase domain (Brecher et al., 2015). Inhibition of MTase will be detrimental to the progression of ZIKV.

The conserved RdRp domain allows for the initiation of RNA synthesis, generating both plus and minus strand RNAs. As with most polymerases, the structure of the enzyme resembles a shape analogous to a right hand with a finger, thumb, and palm region (Papageorgiou et al., 2014; Shanmugam, Velmurugan, & Gromiha, 2016). The human body does not contain an RdRp enzyme or analogs of it, thus inhibitors may not cause severe toxic effects, making it an optimal target in drug design (Shanmugam et al., 2016).

To date, no anti-ZIKV drugs are clinically available, thus, new research methods are being developed with the purpose of identifying target molecules. Recent research has found that ZIKV targets neuronal cells (Millichap, 2016; Miner & Diamond, 2016; Mlakar et al., 2016; Tang et al., 2016). Consequently, any new drugs that may be discovered will have to pass through the blood–brain barrier. Molecular modeling and computational methods are important tools in the development of novel inhibitors of ZIKV (Ekins et al., 2016; Ramharack & Soliman, 2016). A number of inhibitors of the *flavivirus* NS5 protein have been discovered via virtual screening and computational analysis (Brecher et al., 2015; Idrus, Tambunan, & Zubaidi, 2012; Lim & Shi, 2013). Structure-based virtual screening (SBVS) identifies energetically advantageous binding affinities of ligands into a target's active binding site. This allows for new insights on the nature of the active site and the protein–ligand

interactions (Kumalo & Soliman, 2015). The method identifies selective molecules from an extensive library of compounds to dock within a target's active site. Although scoring techniques are used when molecules are docked to the target, literature shows that a large number of final hits are generated, as the compounds docked may be in various geometric poses (Kroemer, 2007). Ligand-based virtual screening generates libraries of compounds based on a known compound or compounds and its illustrative interactions with a particular target (Cele, Muthusamy, & Soliman, 2016).

In an attempt to develop pharmacophore-based modeling, we previously presented a per residue energy decomposition (PRED) protocol where candidates for SBVS were chosen on the position of 3D moieties with an experimentally known compound, thus creating a pharmacophore model based on highly contributing amino acid residues to the bound inhibitor. This approach is based on interactions that occur at a molecular level, including charge, hydrophobic interactions and hydrogen bonding (Cele et al., 2016). The highly contributing residues are identified based on free energy footprints from molecular dynamic (MD) and thermodynamic calculations (Cele et al., 2016; Kumalo & Soliman, 2015; Soliman, 2013). This proves to be an incredibly concise method, rather than 'shooting in the dark' with millions of available small molecules.

In our previous work, we created a possible homology model of the NS5 protein containing both MTase and RdRp domains (Ramharack & Soliman, 2016). Due to the indeterminateness surrounding the ZIKV NS5 protein and potential inhibitors, we will compare our top hits against known inhibitors of the *flavivirus* NS5 protein.

This study will implement the above-mentioned PRED pharmacophore technique in the discovery of potential ZIKV NS5 protein inhibitors, thus aiding medicinal chemists in the synthesis of possible drug candidates.

2. Computational methods

A route map to PRED-based pharmacophore virtual screening approach is depicted in Figure 1.

2.1. Homology modeling and identification of active binding sites of NS5

In our previous article and during the current study, due to the absence of a crystal structure of ZIKV NS5 protein, a homology model was created using the protein sequence obtained from NCBI (Accession number: YP_009227205) (Ramharack & Soliman, 2016). The templates for sequence alignment were identified from

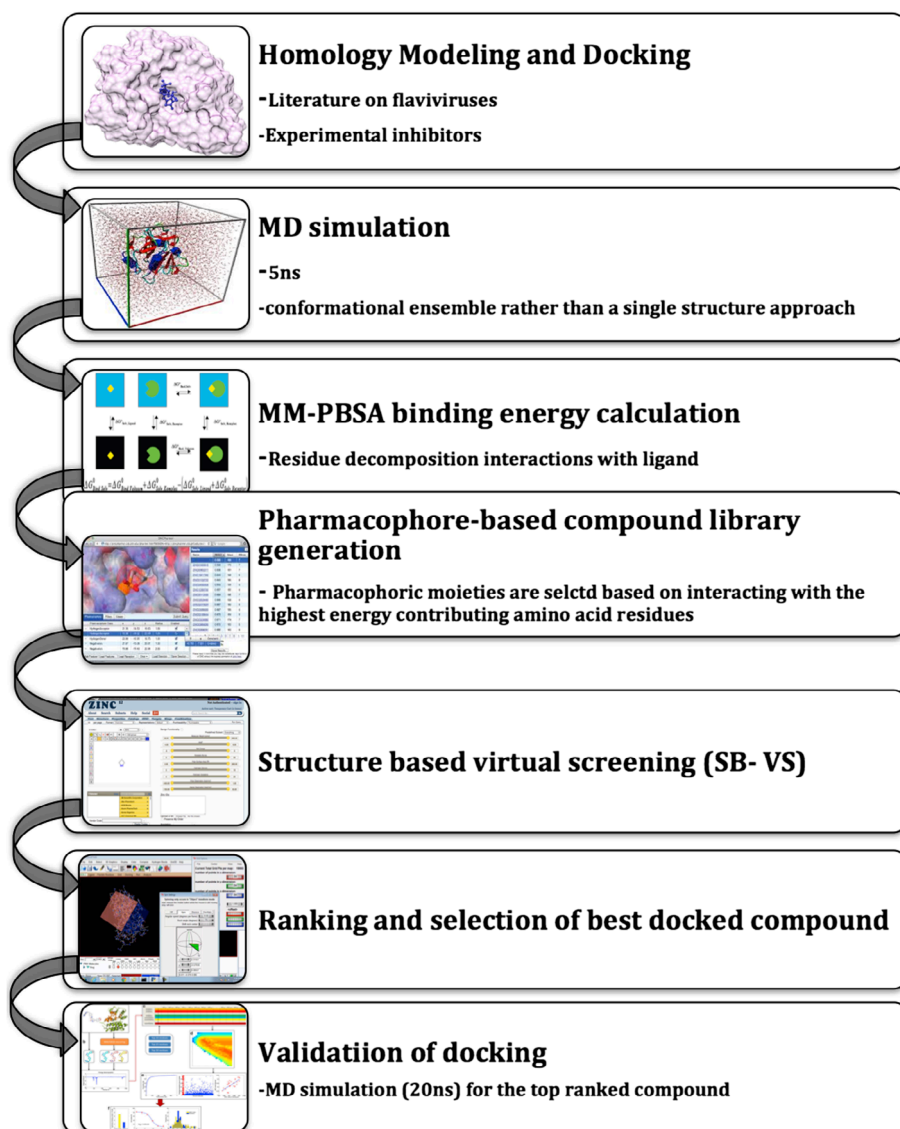


Figure 1. A PRED-based approach outline applied in the study.

Table 1. Criteria summary of chosen templates used in Building the ZIKV NS5 homology model.

Template PDB code	Query cover (%)	Structural identity (%)	E-value
4K6M_A	97	69	0
5CCV_A	98	67	0
2HFZ_A	67	72	0
3VWS_A	68	69	0

NCBI using BLASTp (accessed on 5 March 2016) (Madden, Tatusov, & Zhang, 1996) to find suitable templates, from RCSB protein databank (Berman et al., 2002), for homology modeling. Based on the criteria of identity score, e-value and query cover accuracy (Table 1), the NS5 protein was modeled by using four crystal structures of *flavivirus* enzymes as templates:

Chain A of full-length Japanese Encephalitis Virus NS5 (PDB Code: 4K6M_A); Chain A of full-length NS5 from Dengue virus Type 3 (PDB Code: 5CCV_A); Chain A of RNA Dependent RdRp domain from Nile West Virus (PDB Code: 2HFZ_A); and Chain A of Dengue Serotype 3 RNA-dependent RdRp bound to Nitd-107 (PDB Code: 3VWS_A).

Homology modeling was performed using the Modeller Software version 9.1 (Eswar et al., 2006) add-on in chimera (Yang et al., 2012), in which all three templates were selected to build the model. Multiple sequence alignment was performed using the CLUSTALW server (Sievers et al., 2011), where Chain A of the full-length Japanese Encephalitis Virus NS5 was evidenced to have the best template with the highest identity score (Figure S1). The sequence of the target protein was uploaded to PSIPRED V3.3 (Buchan, Minneci, Nugent, Bryson, & Jones, 2013) in order to obtain a predicted 2D secondary structure of the enzyme. Comparing the homolog to the predicted 2D structure and assessment of the bond angles and torsional strain validated the homology model. A Ramachandran plot for the analyses of bond angles and torsional strain was generated using Maestro (Schrodinger). MolProbity (Chen et al., 2010) results showed 97.2% of all residues were in the favored regions and 99.2% of all residues were in the allowed regions, which left a list of 7 outliers. The active-site residues were determined using Chimera Multi-align Viewer and validated using the SiteHound-web program (Hernandez, Ghersi, & Sanchez, 2009). The list shows that none of the active-site residues are part of these outliers. All results can be found in our previous article and supplementary material (Figure S5) (Ramharack & Soliman, 2016). After completion of the study, the crystal structure of the ZIKV NS5 protein was released. To validate the homology model of the NS5, it was superimposed with the newly released crystal structure (PDB code: 5TFR), showing their structural similarity and validating the model's use for subsequent analysis (Figure 2).

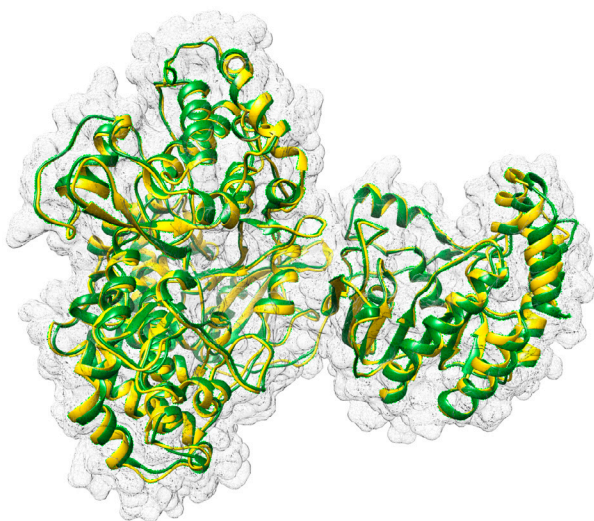


Figure 2. Superimposition of homology model (yellow) with the newly released crystal structure (green-PDB code: 5TFR), showing their structural similarity and validating the model's use for subsequent analysis.

2.2. System preparation

The NS5 modeled structure was separated into two domains, being the MTase of the C-terminal and the RNA-Dependent RdRp of the N-terminal. Experimental drug inhibitors of *flaviviruses* were chosen to dock within each domain's active site.

2.3. Molecular docking of experimental flaviviruses

The small molecules, BG323 and Ribavirin, potent inhibitors of flavivirus NS5, were chosen to dock at the MTase site and the RdRp site, respectively (Leyssen, De Clercq, & Neyts, 2000; Lim & Shi, 2013; Tambunan et al., 2014).

Docking of the compounds was conducted using the AutoDock Vina (Trott & Olson, 2010) software. The procedure was run using the software default settings. The grid box used to define the screening site was elucidated using the AutoDock Vina functionality built into Chimera (Pettersen et al., 2004). The gridbox size and center parameters for the MTase were $x(54, -63.23)$, $y(80, 56.72)$, and $z(54, 10.22)$, respectively, and the RdRp gridbox dimensions were $x(40, -9.69)$, $y(38, 20.41)$, and $z(40, 16.50)$. AutoDock Vina generated results in the pdbqt format and the optimal geometric conformation having the best binding energy was selected from the ViewDock feature and saved in complex with the reference enzyme. The enzyme and ligand for each system was prepared using Chimera (Yang et al., 2012) and MMV molecular modeling suites (Kusumaningrum, Budianto, Kosela, Sumaryono, & Juniarti, 2014) and subsequently subjected to MD simulations.

2.4. Molecular dynamic (MD) simulations

The MD simulation was performed using the GPU version of the PMEMD engine provided with the Amber 14 package. The FF14SB force field of the Amber package (Nair & Miners, 2014) was used to describe the complex.

ANTECHAMBER (Wang, Wang, Kollman, & Case, 2006) was used to generate atomic partial charges for the ligands by utilizing the Restrained Electrostatic Potential (RESP) and the General Amber Force Field (GAFF) procedures. The Leap module of Amber 14 allowed for addition of hydrogen atoms to the systems as well as Na^+ and Cl^- counter ions for neutralization.

The system was suspended implicitly within an orthorhombic box of TIP3P water molecules such that all atoms were within 8 Å of any box edge.

An initial minimization of 2000 steps was carried out with an applied restraint potential of 500 kcal/mol Å² for both complexes. An additional full minimization of 1000 steps was further carried out by conjugate gradients algorithm without restraint.

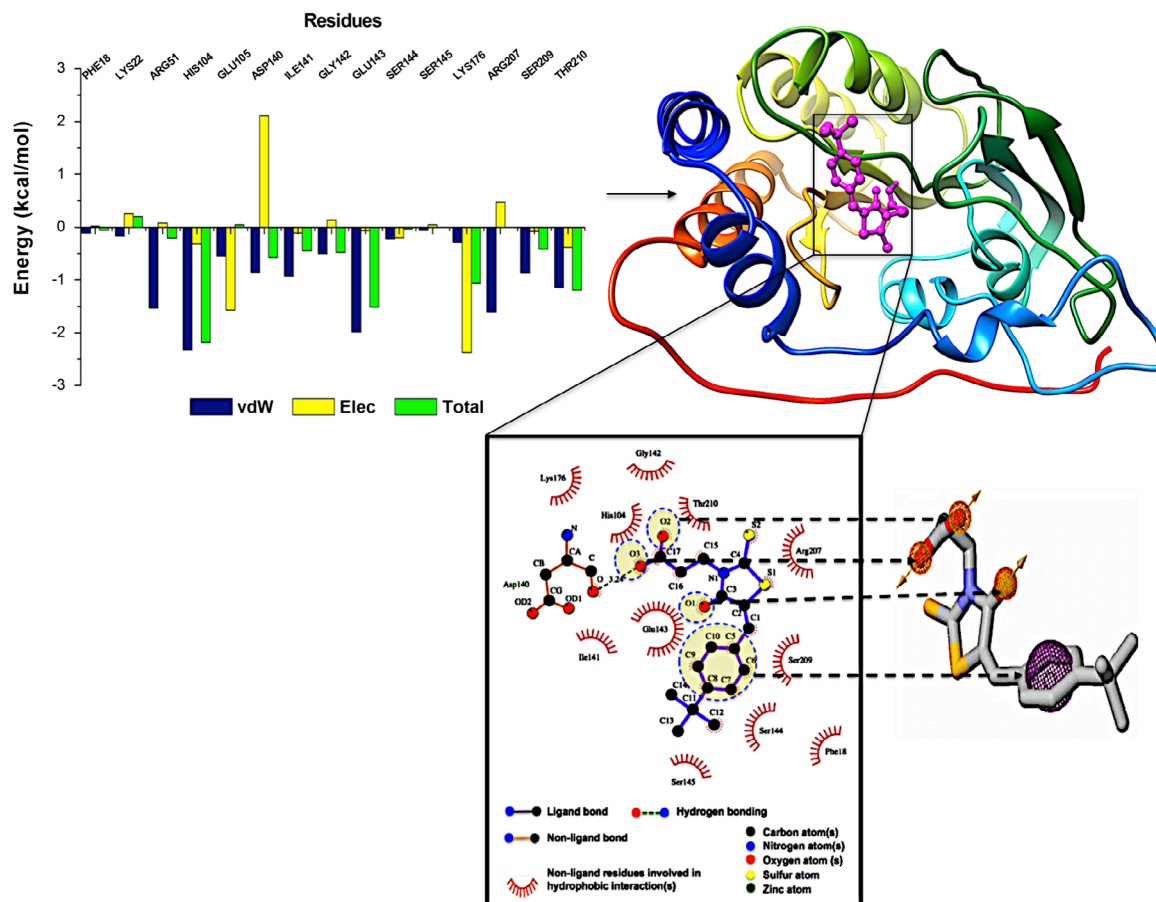


Figure 3. The steps taken toward creating the pharmacophore model from the MTase-BG323 complex. The yellow circles spotlight the pharmacophoric moieties that were chosen for the model, based on the highest contributing residues, depicted in the binding affinity graph.

A gradual heating MD simulation from 0–300 K was executed for 50 ps, such that the system maintained a fixed number of atoms and fixed volume, i.e. a canonical ensemble (NVT). The solutes within the system are imposed with a potential harmonic restraint of 10 kcal/mol Å² and collision frequency of 1.0 ps⁻¹. Following heating, an equilibration estimating 500 ps of the each system was conducted, the operating temperature was kept constant at 300 K. Additional features such as a number of atoms and pressure were also kept constant mimicking an isobaric–isothermal ensemble (NPT). The systems pressure was maintained at 1 bar using the Berendsen barostat.

The total time for the MD simulation conducted was 5 ns. In each simulation the SHAKE algorithm was employed to constrict the bonds of hydrogen atoms. The time step of each simulation was 2 fs and an SPFP precision model was used. The simulations coincided with isobaric–isothermal ensemble (NPT), with randomized

seeding, constant pressure of 1 bar maintained by the Berendsen barostat, a pressure-coupling constant of 2 ps, a temperature of 300 K, and Langevin thermostat with collision frequency of 1.0 ps⁻².

Coordinates were saved every 1 ps and the trajectories were analyzed every 1 ps using the PTRAJ module employed in Amber14.

2.5. Binding free energy calculations

To estimate the binding affinities of each system, the binding free energies were calculated using the molecular mechanics/GB surface area method (MM/GBSA) (Genheden & Ryde, 2015). Binding free energies were averaged over 5000 snapshots extracted from the 5 ns trajectory. The free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

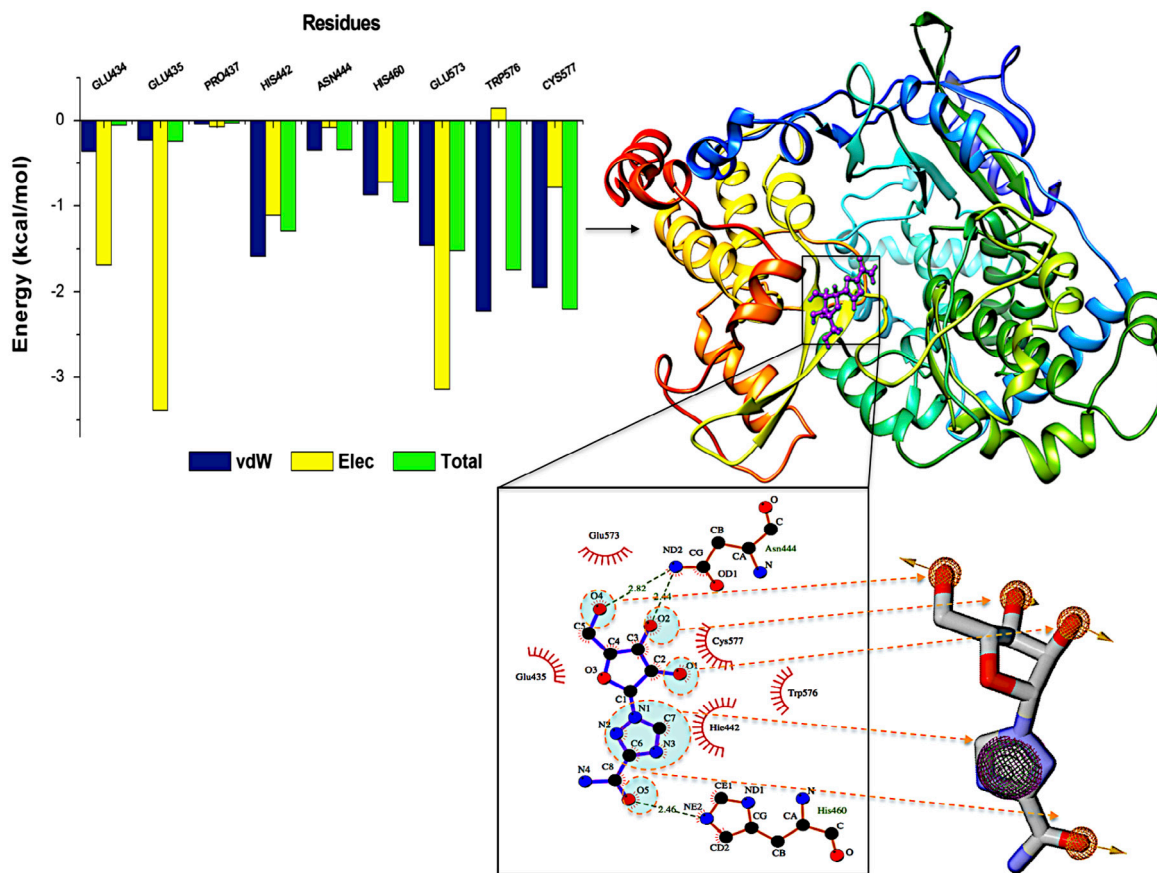


Figure 4. RdRp–Ribavirin complex ligplot analysis – creating the pharmacophore model to virtually screen for new RdRp potential lead compounds. The yellow circles spotlight the pharmacophoric moieties that were chosen for the model based on the highest contributing residues, depicted in the binding affinity graph.

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

$$\Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (4)$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (5)$$

The term E_{gas} denotes the gas-phase energy, which consist of the internal energy E_{int} , Coulomb energy E_{ele} , and the van der Waals energies E_{vdw} . The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol} , was estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G . The non-polar solvation energy, SA.

G_{SA} , was determined from the solvent accessible surface area, using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB} , contribution was estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature, respectively.

To obtain the contribution of each residue to the total binding free energy profile between the inhibitors Ribavirin and BG323 with RdRp and MTase, respectively, per-residue free energy decomposition was carried out at the atomic level for imperative residues using the MM/GBSA method in Amber 14.

2.6. Pharmacophore model creation and library generation

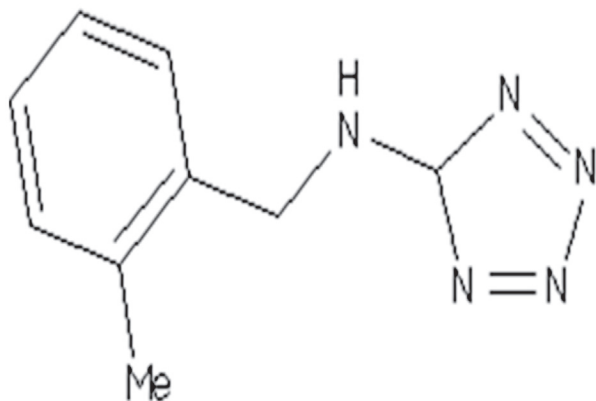
The inhibitors Ribavirin and BG323 were first simulated at the active site of RdRp and MTase respectively, for 5 ns, to create the bound conformation of the ligands. Both these compounds have experimentally exhibited ZIKV inhibition in *in vitro* and *in vivo* models (Mumtaz, van Kampen, Reusken, Boucher, & Koopmans, 2016; Sweeney et al., 2015; Zmurko et al., 2016). Per-residue energy decomposition analysis was used to determine the amino acids that contribute the most towards ligand binding. The pharmacophoric moieties that interacted with the highly contributing residues were then chosen

Table 2. Representation of the top three compounds bound to MTase and RdRp. The compounds, ZINC64717952 and ZINC39563464 showed the best binding affinity to MTase and RdRp, respectively.

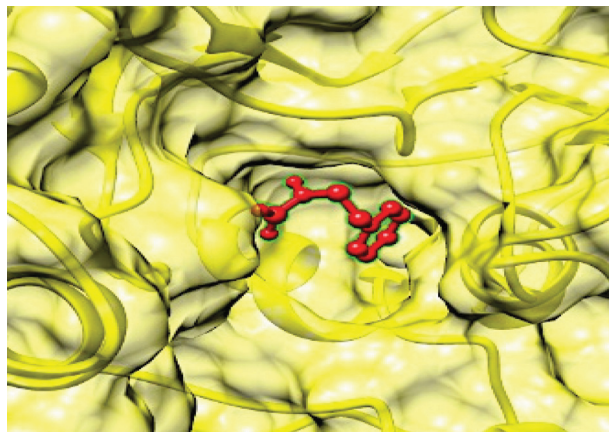
MTase

Compound

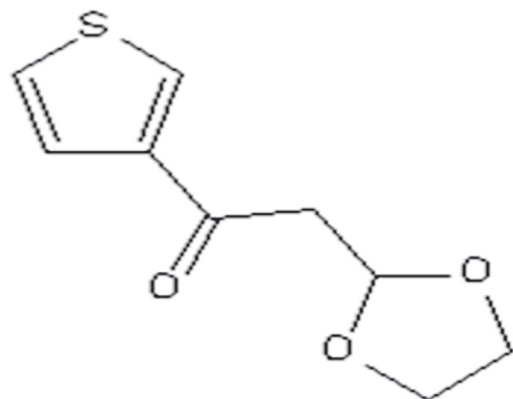
Docked structure



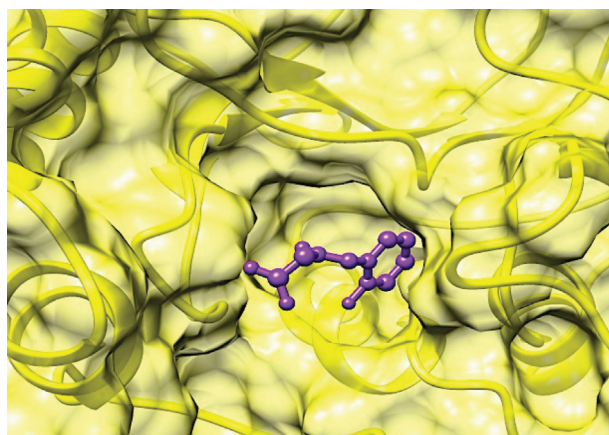
ZINC64717952



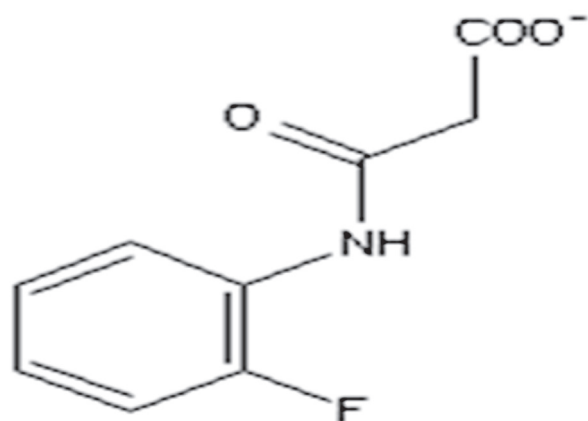
$\Delta G = -6.1$ kcal/mol



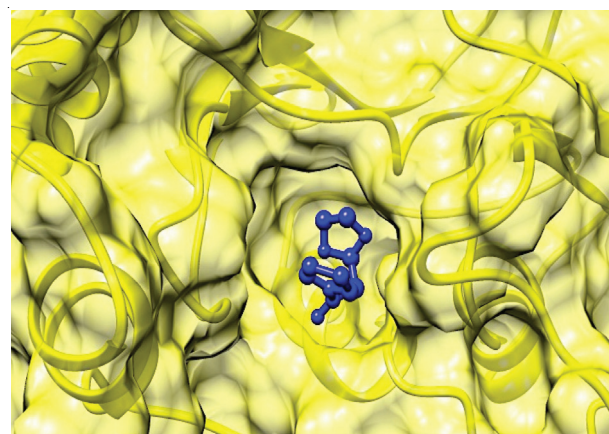
ZINC85652269



$\Delta G = -5.4$ kcal/mol



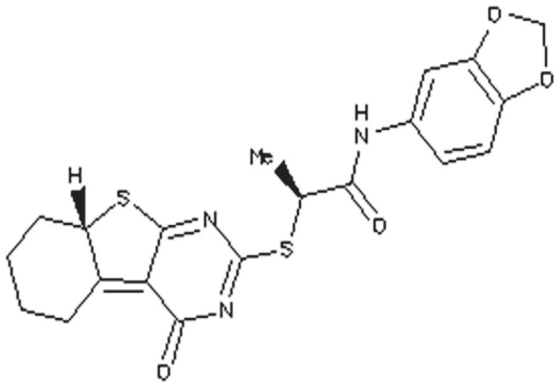
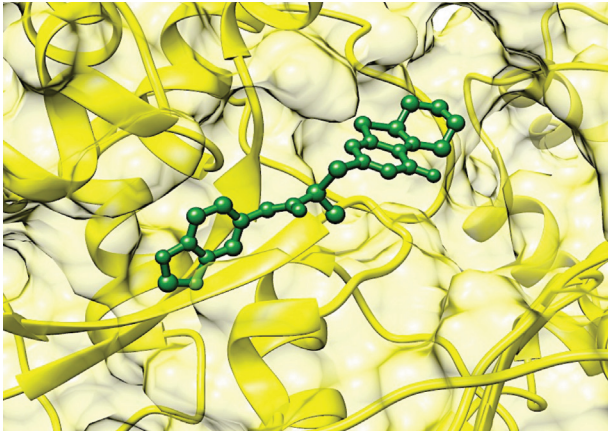
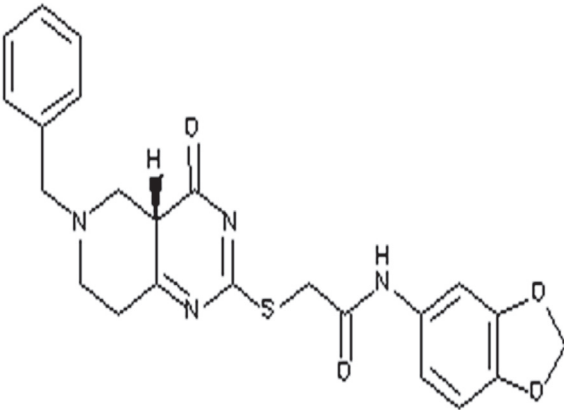
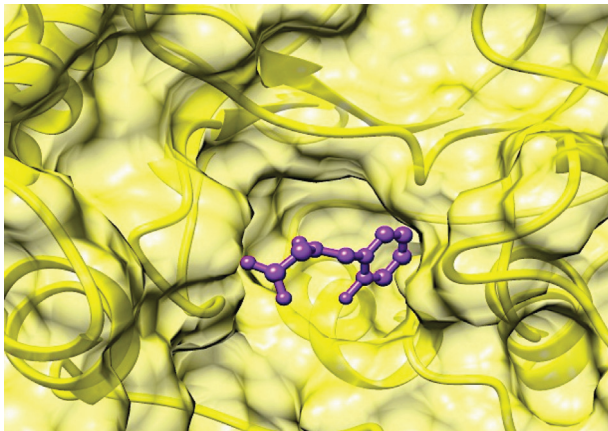
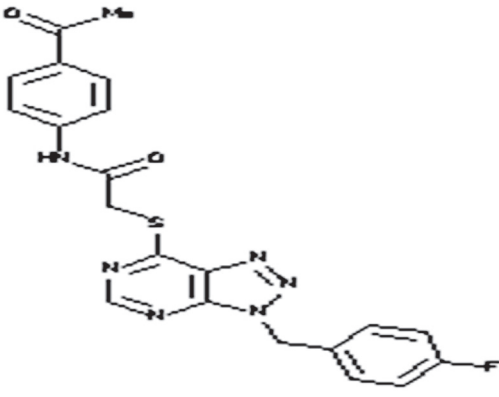
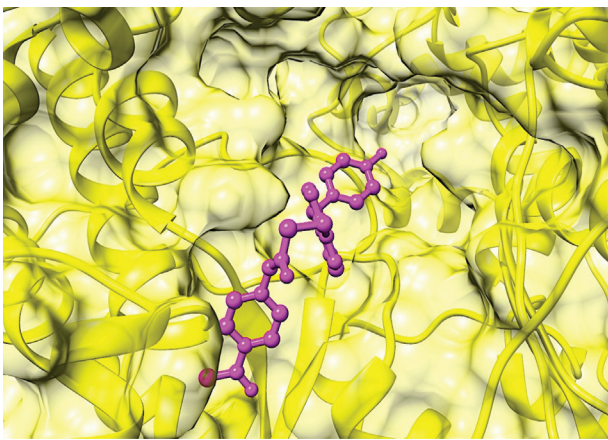
ZINC09304524



$\Delta G = -5.1$ kcal/mol

(Continued)

Table 2. (Continued).

RdRp	
Compound	Docked structure
 ZINC39563464	 $\Delta G = -9.1$ kcal/mol
 ZINC39588253	 $\Delta G = -7.9$ kcal/mol
 ZINC11758496	 $\Delta G = -7.5$ kcal/mol

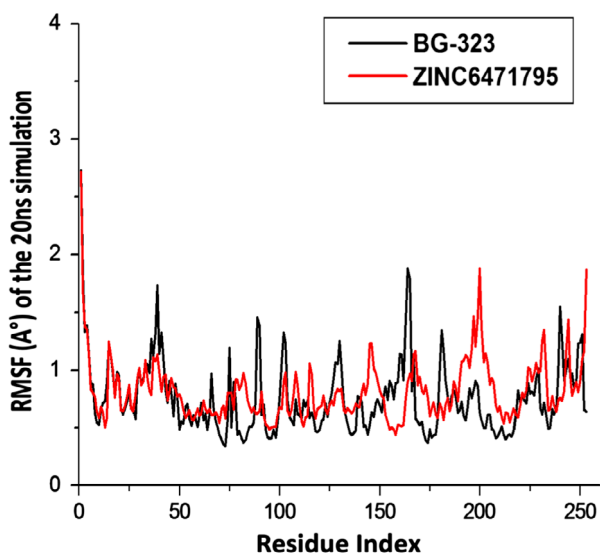


Figure 5. The $C\alpha$ RMSF of MTase-BG323 and MTase-ZINC64717952 during the molecular dynamic simulation.

to construct our model. The model was then added to ZincPharmer (Koes & Camacho, 2012), with specific selection criteria (molecular weight of <500 Da, rotatable bonds <6, hydrogen bond donors <5, and hydrogen bond acceptors <10), to screen the ZINC database (Irwin & Shoichet, 2005). Lipinski's rule of five and toxicity (ADMET) properties were used as filters to remove non-drug-like hits (Lipinski, Lombardo, Dominy, & Feeney, 2012).

2.7. Structure-based virtual screening

The drug-like hits identified using our protocol were subjected to SBVS. Docking was carried out to differentiate between ligands based on the molecules' geometric characteristics that allow it to bind to the enzyme's active site (Forli et al., 2016). The docking calculations were performed using Autodock Vina (Trott & Olson, 2010). During docking, Gasteiger partial chargers were assigned and the Autodock atom types were defined using the Autodock Graphical user interface supplied by MGL tools (Sanner, Olson, & Spehner, 1996). The docked conformations were generated using the Lamarckian Genetic Algorithm (Morris & Huey, 2009). The Raccoon software was used to convert the files into a compatible pdbqt format required for docking. The gridbox was

defined using Autodock Vina. The calculation reports for each ligand conformation in its respective complex were analyzed to obtain affinity energy (kcal/mol). During the docking process, a maximum of 50 conformers was considered for each compound. After screening, molecular docking and filtering, the ligand with the highest affinity towards the agonist was selected from the library.

2.8. Validation of docking approach

Previous experiences have verified that docking may result in the best geometric conformation of the docked complex, however, short MD simulations may not be able to maintain the stability of the complex and thus lead to the molecules being disorientated. Thus, to validate the approach applied in this study, the most favorable Mtase and RdRp complex was subjected to further MDs studies (20 ns). The procedure for MDs simulation was the same as in "MD simulations" section and thermodynamic calculations as "Binding Free Energy Calculations" section.

2.9. Assessment of drug likeness

The online software SwissADME was used to compute the physicochemical descriptors as well as predict the pharmacokinetic properties and drug-like nature of the screened compounds compared to that of BG323 and Ribavirin. to (Bultet et al., 2016; Daina, Michielin, & Zoete, 2014). SwissADME utilizes the 'Brain Or Intestinal Estimated permeation (BOILED-Egg)' method which computes the lipophilicity and polarity of small molecules (Daina & Zoete, 2016).

3. Results and discussion

3.1. Homology model and active binding site determination

Due to the absence of a crystal structure for the Zika NS5 enzyme, a homology model, having a zDope score of -0.76 was generated, and validated using a ramachandran plot. The active site residues were determined for both the MTase and RdRp region (Figure S1–S3). The comprehensive set of results are presented in our previous publication (Ramharack & Soliman, 2016). To further validate both the MTase and RdRp, the homology model was superimposed to the newly released crystal structure of the Zika NS5 (Figure 2), using Chimera (Pettersen et al., 2004).

Table 3. The comparison of MTase's binding affinity with BG323 and ZINC64717952.

Compound	Energy components (kcal/mol)				
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
BG323	-33.32 ± 1.82	-11.84 ± 1.68	-49.16 ± 1.86	20.83 ± 1.80	-28.33 ± 1.87
ZINC64717952	-35.77 ± 2.66	-10.47 ± 2.68	-46.24 ± 3.97	19.74 ± 2.14	-26.50 ± 3.14

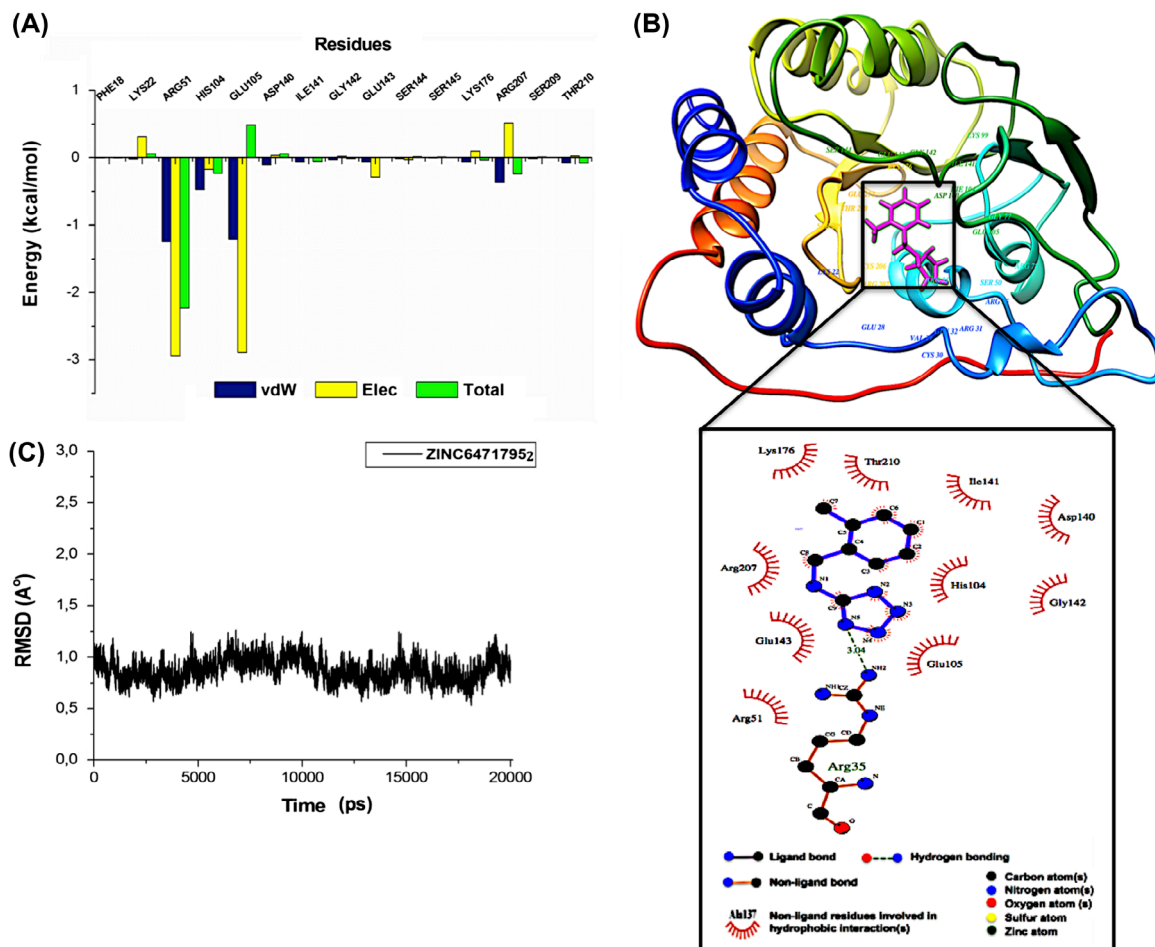


Figure 6. MTase–ZINC64717952 complex interactions (A) Per-residue decomposition analysis showing Arg51 and Glu105 to have the greatest bond fluctuations (B) Ligplot depiction of hydrophobic and hydrogen bond interactions in the complex which was validated by (C) The time evolution of RMSD of the C-alpha atom backbone of the MTase–ZINC64717952 complex.

3.2. PRED pharmacophore model

In this study, a pharmacophore hypothesis was adopted by utilizing per residue decomposition energy-based approach. The structural features of a protein as well as the chemical characteristics of a ligand are employed in the construction of a pharmacophore model. To generate the pharmacophore model, a 5 ns MD simulation was run on complexes (MTase–BG323) and (RdRp–Ribavirin), followed by PRED computed from MM/GBSA calculations. The MM/GBSA approach has proven to be, in principle, accurate in both scoring function and binding free energy results (Genheden & Ryde, 2015; Hayes, 2009). This allows for improved pharmacophore modeling and thus the generation of a concise library of small molecules. The MTase–BG323 complex showed His104 (−2.176 kcal/mol), Glu143 (−1.846 kcal/mol), Thr210 (−1.192 kcal/mol), and

Lys176 (−1.061 kcal/mol) to be the highest contributing residues to interact with the ligand. Strong hydrophobic interactions were formed between Glu143 and the benzene ring of BG323, while, energetically favorable residue, Asp140, formed hydrogen bonds with the terminal hydroxyl groups of the ligand (Figure 3).

Hydrogen bonds between the RdRp–Ribavirin complex included residues Asn444 (−1.296 kcal/mol) and His460 (−0.956 kcal/mol), while the contributing residues; namely, GLU573 (−1.521 kcal/mol), TRP576 (−1.744 kcal/mol) and Cys577 (−2.202 kcal/mol) were involved in hydrophobic interactions with the ligand. The features from each complex were used as a query on ZINCpharmer (Koes & Camacho, 2012) to create the PRED-based pharmacophore (Figure 4). Results revealed 18 hits obtained from the MTase–BG323 pharmacophore and 23 hits from the RdRp–Ribavirin pharmacophore.

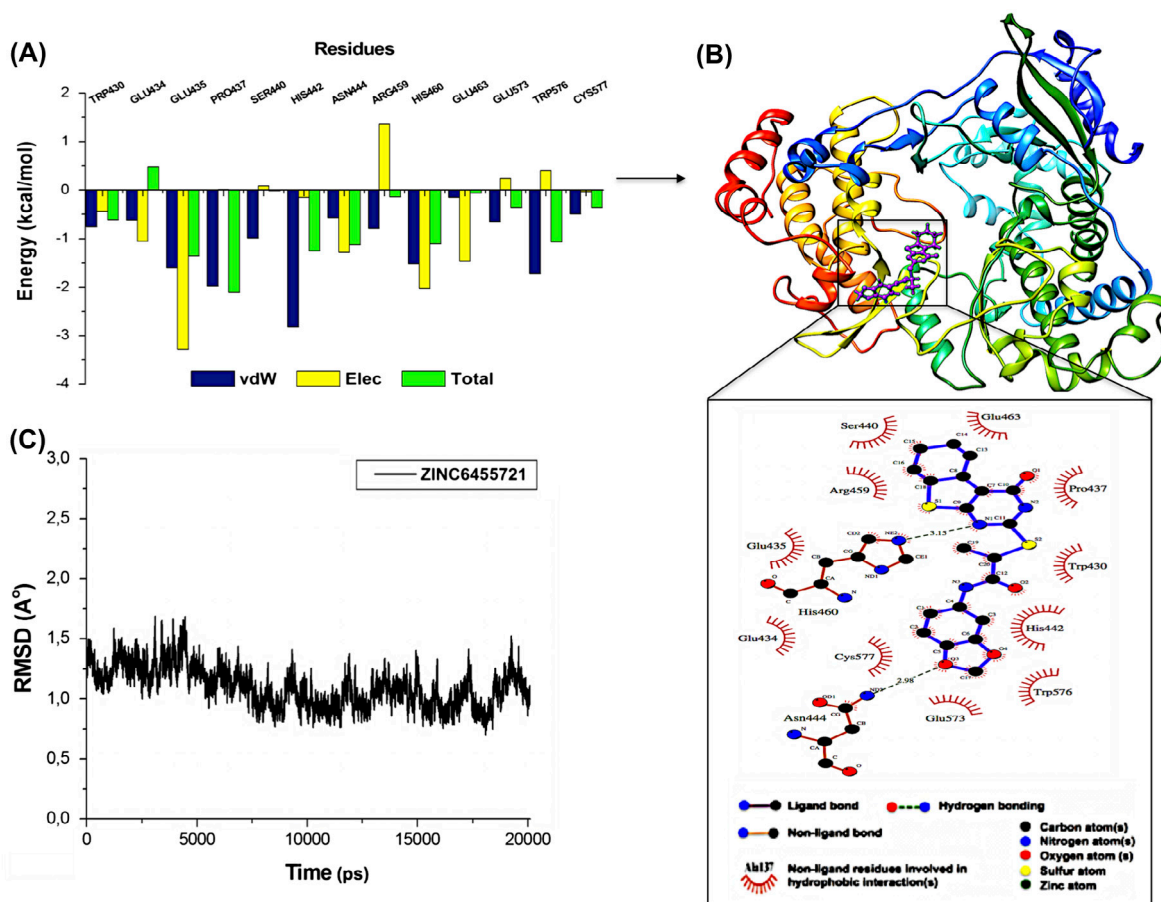


Figure 7. RdRp–ZINC39563464 complex interactions (A) Per-residue decomposition analysis showing Arg459 and Glu435 to have the greatest bond fluctuations, (B) Ligplot depiction of hydrophobic and hydrogen bond interactions in the complex which was validated by (C) The time evolution of RMSD of the C-alpha atom backbone of the MTase–ZINC39563464 complex.

Table 4. The comparison of RdRp's binding affinity with Ribavirin and ZINC39563464.

Compound	Energy components (kcal/mol)				
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
Ribavirin	-23.20 ± 3.13	-40.92 ± 13.03	-64.12 ± 13.34	47.59 ± 9.93	-16.53 ± 4.84
ZINC39563464	-38.17 ± 5.39	-17.32 ± 5.99	-55.49 ± 8.85	30.45 ± 4.85	-25.04 ± 5.35

3.3. Molecular docking

To further refine and reduce false positives retrieved from the hit compounds, the hits for each complex were subjected to molecular docking within the active sites of MTase (18 hits) and RdRp (23 hits). This assessed their geometric feasibility at each domain, leading to only three top ranked compounds (Table 2). Based on the interactions and binding affinities of the respective three top-ranked compounds to MTase and RdRp, ZINC64717952 and ZINC39563464 were chosen as respective top hits. Each complex was subsequently subjected to MD studies to elucidate on the enzyme–ligand

interactions of the two potential inhibitors under virtual conditions.

3.4. MD simulations and binding free energy analysis

The MTase–ZINC64717952 and RdRp–ZINC39563464 complexes were subjected to a 20 ns MD simulation in order to check the convergence dynamic stability and to analyze the energetics of each complex. The RMSD profiles of the MTase-complex and RdRp-complex indicate that both systems were stable during the simulation (Figures 6(C) and 7(C)).

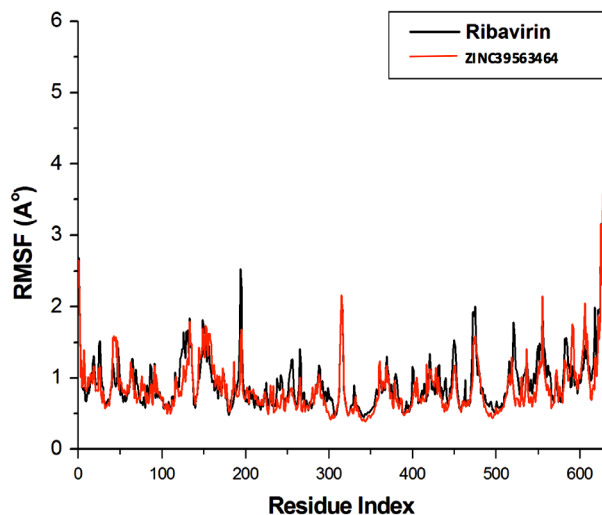


Figure 8. The stable Ca RMSF of RdRp–Ribavirin and RdRp–ZINC39563464 during the molecular dynamic simulation.

3.4.1. MTase–ZINC64717952 Complex

The docked MTase–ZINC64717952 complex showed ionic interactions involving seven residues common to MTase–BG323 (Glu143, Arg207, Lys176, Thr210, Ile141, Asp140, and Gly142). Interestingly, however, a hydrogen bond was noted between the nitrogen of Arg35 and the aromatic ring of ZINC64717952, this was peculiar, as Arg35 was not involved in any ionic interactions of the MTase–BG323 complex. The MTase–ZINC64717952 complex used Asp140 as a hydrogen bond acceptor, whereas, the MTase–BG323 complex depicted hydrophobic interactions between Asp140 and the benzene ring of BG323. These ionic bond deviations between systems may be due to the size of ZINC64717952 in comparison to BG323. ZINC64717952 was significantly reduced in size, containing predominantly the heterocyclic rings from the pharmacophore model. Due to the size of ZINC64717952, the nitrogen of aromatic ring was allowed to form a hydrogen bond with the amine group (Arg35) further into the hydrophobic pocket of MTase. Docking results showed the same binding affinity in both complexes, however, receptor residue stability showed increased fluctuations in the ZINC64717952–MTase complex compared to the experimental complex (Figure 5). The overall compactness of the receptor was measured by the radius of gyration (around the Ca atoms) and was indicative of greater fluctuations of the MTase–ZINC64717952 complex compared to the experimental complex (Figure S6), verifying the root mean square fluctuations (RMSF) seen in Figure 5. Although ZINC64717952 docked in a structurally favorable manner, MM/GBSA analysis showed free binding energy of

the MTase–BG323 complex (-28.70 kcal/mol) to be higher in magnitude than that of MTase–ZINC64717952 (-26.50 kcal/mol).

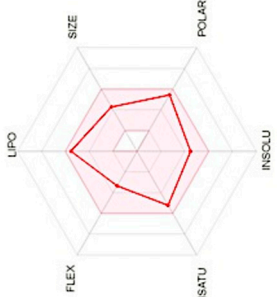
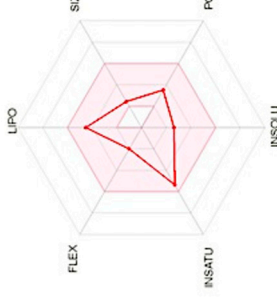
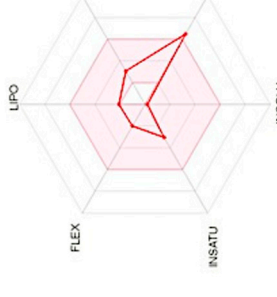
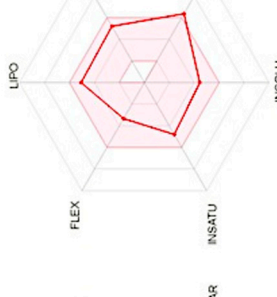
The tetrazole aromatic ring in ZINC64717952 contains highly active nitrogen atoms, increasing electronegativity and steric hindrance (Ostrovskii, Trifonov, & Popova, 2012). The Generalized Born (GB) method is used to calculate the molecular electrostatic forces in solvent. Table 3 shows ZINC64717952 to have elevated columbic energy, thus leading to increased gas-phase energy, validating the free energy analysis (Figure 6(A)) (Genheden & Ryde, 2015). This, however, does not rule-out the possibility of ZINC64717952 as a potential inhibitor of the MTase enzyme as the intermolecular forces between the receptor and ligand were favorable. This study will have to be evaluated *in vitro*, where the further analysis may reveal the inhibitory potential of the compound.

3.4.2. RdRp–ZINC39563464 complex

The docked RdRp–ZINC39563464 complex showed ZINC39563464 to interact with nitrogen atoms of two residues; Asn444 and His460. Notably, the nitrogen atoms from the same residues form hydrogen bonds with the terminal oxygen of the Ribavirin, showing consistent residue interactions of the experimental ligand and ZINC39563464. These hydrogen interactions are formed from non-covalent bonding of the hydrogen donor (Asn444 and His460) with the acceptors (oxygen and nitrogen) of the ligand. This articulates the directionality and specificity of the active site's β -strand recognition of both Ribavirin and ZINC39563464. The complex exhibiting a relatively stable RMSD profile during the simulation further validated this (Figure 7(C)). The pharmacophoric hot spot residue, His442, formed hydrophobic bonds with the aromatic rings of both Ribavirin and ZINC39563464. It is noteworthy that four other hydrophobic-interacting residues; Cys577, Trp576, Glu573 and Glu435 were common to both ligands, thus stabilizing both energetically favorable ligands in the available hydrophobic pocket. Table 4 depicts the analysis of binding free energy by the use of MM/GBSA of the RdRp–ZINC39563464 complex was used to support the docking results.

The predicted binding free energy for the complex was -25.04 kcal/mol, which is considerably higher in magnitude than that achieved by the RdRp–Ribavirin complex (-16.53 kcal/mol), thus confirming the docking results and indicating a stronger binding of ZINC39563464 to RdRp compared to the experimental ligand (Figure 7(A) and (B)). The relatively large size of the ligand could explain the increased number of residues encompassing apparent hydrophobic interactions with ZINC39563464, and could substantiate the exhibition of stable RdRp residues by RMSF profiling

Table 5. The comparison of drug likeness of the screened compounds compared to that of the experimental drugs against ZIKV.

	BG323	ZINC6471795	Ribavirin	ZINC39563464
Molecular formula	C17H19NO3S2	C9H11N5	C8H12N4O5	C20H19N3O4S2
Molecular weight g/mol	349.47	189.22	244.20	429.51
Lipophilicity (iLOGP)	2.79	2.47	0.22	2.86
Water soluble	Moderately soluble	Soluble	Soluble	Moderately Soluble
GIT absorption	High	High	Low	Low
BBB permeability	Not permeable	Semi-permeable	Not permeable	Semi-Permeable
Bioavailability score	0.56	0.55	0.55	0.55
Synthetic accessibility	3.39	2.93	3.89	4.89
Druglikeness (Lipinski)	Yes	Yes	Yes	Yes
‘Boiled-Egg’ method summary				

(Figure 8). As an additional check, the radius of gyration (Rg of the Ca atoms) was compared in both simulations to provide a measure of overall compactness of the protein (Figure S6). The fluctuations of Rg stayed with 1 Å in both simulations indicative of a stable protein complex with both experimental and screened compound.

3.4. Assessment of drug likeness

Ribavirin has a plethora of side effects including thrombocytopenia, myalgia, leucopenia, and cognitive impairment (Hinton et al., 2016; Kryger, Wohl, Smith, & Zelikin, 2013; Munir et al., 2010). This proves to be a challenge when trying to inhibit a virus that already causing these symptoms. BG323 is a new compound that has been proven to have potent effects on *flavivirus* NS5 proteins, however, the compound is unable to pass the blood–brain barrier, making it difficult to act on ZIKV-targeted neuronal cells (Miner & Diamond, 2016; Tang et al., 2016). The possible pro-drugs of ribavirin and BG323, being, ZINC39563464 and ZINC64717952, respectively, can be described as potential lead compounds after assessment through SwissADME (Table 5) (Bultet et al., 2016).

4. Conclusion

ZIKV is a rapidly evolving virus that has had detrimental long-term effects over a very short period of time. This study proposes two new compounds that have shown promising physicochemical properties and strong interactions with ZIKV MTase and RdRp, thus validating the PRED model as an effective strategy to enhance typical virtual screening methods for the rapid identification of potential lead compounds as inhibitors against pathogenic biological targets such as ZIKV. This strategic *in silico* technique will serve as a beneficial tool to enhance drug discovery and decrease excessive wastage of financial and experimental resources by synthesizing large numbers of compounds that may not be beneficial in the inhibition of target enzymes. The lead compounds, ZINC64717952 and ZINC39563464, have shown substantial stability in complex with the target enzymes and thus further experimental analysis is necessary for efficacy and toxicity validation.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Research Foundation [grant number 102103]; the College of Health Sciences, UKZN and Center for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

Supplementary material

The supplementary material for this paper is available online at <http://dx.doi.org/10.1080/07391102.2017.1313175>.

References

- Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., ... Westbrook, J. D. (2002). The Protein Data Bank. *Acta Crystallographica Section D Biological Crystallography*, 58, 899–907.
- Brecher, M., Chen, H., Liu, B., Banavali, N. K., Jones, S. A., Zhang, J., ... Li, H. (2015). Novel broad spectrum inhibitors targeting the flavivirus methyltransferase. *PLoS ONE*, 10(6), 1–15. doi:10.1371/journal.pone.0130062
- Brito, C. (2016). Zika virus: A new chapter in the history of medicine. *The Lancet Neurology*, 28, 679–680. doi:10.1016/S1474-4422(16)00074-0
- Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, 41, W349–W357. doi:10.1093/nar/gkt381
- Bultet, L. A., Aguilar-Rodríguez, J., Ahrens, C. H., Ahmé, E. L., Ai, N., Aimo, L., ... Zollinger, A. (2016). The SIB Swiss Institute of bioinformatics' resources: Focus on curated databases. *Nucleic Acids Research*, 44(D1), D27–D37. doi:10.1093/nar/gkv1310
- Campos, G. S., Bandeira, A. C., & Sardi, S. I. (2015). Zika virus outbreak, Bahia, Brazil. *Emerging Infectious Diseases*, 21, 1885–1886. doi:10.32301/eid2110.150847
- Cele, F. N., Muthusamy, R., & Soliman, M. E. S. (2016). Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery: A study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. *Drug Design, Development and Therapy*, 10, 1365–1377.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., ... Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, 66, 12–21. doi:10.1107/S0907444909042073
- Chibueze, E. C., Tirado, V., Lopes, K. da S., Balogun, O. O., Takemoto, Y., Swa, T., ... Oladapo, O. T. (2017). Zika virus infection in pregnancy: A systematic review of disease course and complications. *Bulletin of the World Health Organization*, 6, 1–35. doi:10.2471/BLT.16.178426
- Daina, A., Michielin, O., & Zoete, V. (2014). iLOGP: A simple, robust, and efficient description of n-octanol/water partition coefficient for drug design using the GB/SA approach. *Journal of Chemical Information and Modeling*, 54, 3284–3301. doi:10.1021/ci500467k
- Daina, A., & Zoete, V. (2016). A BOILED-egg to predict gastrointestinal absorption and brain penetration of small molecules. *ChemMedChem*, 11, 1117–1121. doi:10.1002/cmdc.201600182
- Egloff, M. P., Benarroch, D., Selisko, B., Romette, J. L., & Canard, B. (2002). An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: Crystal structure and functional characterization. *European Molecular Biology Organization Journal*, 21, 2757–2768. doi:10.1093/emboj/21.11.2757
- Ekins, S., Mietchen, D., Coffee, M., Stratton, T., Freundlich, J., Freitas-Junior, L., ... Andrade, C. (2016). Open drug discovery for the Zika virus. *Future Research*, 5, 1–14. <http://doi.org/10.12688/f1000research.8013.1>

- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., ... Sali, A. (2006). Comparative protein structure modeling using Modeller. *Current Protocols in Protein Science*, 50, 1–31. doi:10.1002/0471250953.bi0506s15
- Faye, O., Freire, C. C. M., Iamarino, A., Faye, O., de Oliveira, J. V. C., Diallo, M., ... Sall, A. A. (2014). Molecular evolution of Zika virus during its emergence in the 20th century. *PLoS Neglected Tropical Diseases*, 8(1), 1–10. doi:10.1371/journal.pntd.0002636
- Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols*, 11, 905–919. doi:10.1038/nprot.2016.051
- Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, 10, 449–461. doi:10.1517/17460441.2015.1032936
- Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., ... Weaver, S. C. (2012). Genetic characterization of zika virus strains: Geographic expansion of the asian lineage. *PLoS Neglected Tropical Diseases*, 6(2), 1–7. doi:10.1371/journal.pntd.0001477
- Hayes, E. B. (2009). Zika virus outside Africa. *Emerging Infectious Diseases*, 15, 1347–1350. doi:10.3201/eid1509.090442
- Hernandez, M., Ghersi, D., & Sanchez, R. (2009). SITE-HOUND-web: A server for ligand binding site identification in protein structures. *Nucleic Acids Research*, 37(Web Server), W413–W416. doi:10.1093/nar/gkp281.
- Hinton, T. M., Zuwala, K., Deffrasnes, C., Todd, S., Shi, S., Marsh, G. A., ... Zelikin, A. N. (2016). Polyanionic macromolecular prodrugs of ribavirin: Antiviral agents with a broad spectrum of activity. *Advanced Healthcare Materials*, 5, 534–540. doi:10.1002/adhm.201500841
- Idrus, S., Tambunan, U., & Zubaidi, A. A. (2012). Designing cyclopentapeptide inhibitor as potential antiviral drug for dengue virus ns5 methyltransferase. *Bioinformation*, 8, 348–352. doi:10.6026/97320630008348
- Irwin, J. J., & Shoichet, B. K. (2005). ZINC – A free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*, 45, 177–182. doi:10.1021/ci049714+
- Koes, D. R. C., & Camacho, C. J. (2012). ZINCPharmer: Pharmacophore search of the ZINC database. *Nucleic Acids Research*, 40(W1), W409–W414. doi:10.1093/nar/gks378
- Kroemer, R. T. (2007). Structure-based drug design: Docking and scoring. *Current Protein & Peptide Science*, 8, 312–328. doi:10.2174/138920307781369382
- Kryger, M. B. L., Wohl, B. M., Smith, A. A. A., & Zelikin, A. N. (2013). Macromolecular prodrugs of ribavirin combat side effects and toxicity with no loss of activity of the drug. *Chemical Communications (Cambridge, England)*, 49, 2643–2645. doi:10.1039/c3cc00315a
- Kumalo, H. M., & Soliman, M. E. (2015). Per-residue energy footprints-based pharmacophore modeling as an enhanced in silico approach in drug discovery: A case study on the identification of novel β -secretase1 (BACE1) inhibitors as anti-Alzheimer agents. *Cellular and Molecular Bioengineering*, 9, 175–189. doi:10.1007/s12195-015-0421-8
- Kusumaningrum, S., Budianto, E., Kosela, S., Sumaryono, W., & Juniarti, F. (2014). The molecular docking of 1,4-naphthoquinone derivatives as inhibitors of Polo-like kinase 1 using Molegro Virtual Docker. *Journal of Applied Pharmaceutical Science*, 4, 47–53. doi:10.7324/JAPS.2014.4119
- Leyssen, P., De Clercq, E., & Neyts, J. (2000). Perspectives for the treatment of infections with flaviviridae. *Clinical Microbiology Reviews*, 13, 67–82. doi:10.1128/CMR.13.1.67-82.2000
- Lim, S., & Shi, P. (2013). West Nile virus drug discovery. *Viruses*, 5, 2977–3006. doi:10.3390/v5122977
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2012). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, 64, 4–17.
- Lissauer, D., Smit, E., & Kilby, M. D. (2016). Zika virus and pregnancy. *An International Journal of Obstetrics and Gynaecology*, 123, 1258–1263. doi:10.1111/1471-0528.14071
- Madden, T. L., Tatusov, R. L., & Zhang, J. (1996). Application of network BLAST server. *Methods in Enzymology*, 266, 131–141.
- Mahfuz, M., Khan, A., Mahmud, H., Hasan, M., Parvin, A., Rahman, N., & Rahman, S. M. B. (2014). Indian Journal of Pharmaceutical and Biological Research (IJPBR) in silico modeling and immunoinformatics probing disclose the epitope based peptide vaccine against Zika virus envelope glycoprotein. *Indian Journal of Pharmaceutical Biological Research*, 2, 44–57.
- Mahfuz, M., Khan, A., Mahmud, H., Hasan, M., Parvin, A., Rahman, N., & Rahman, S. M. B. (2015). Design and prediction of potential RNAi (siRNA) molecules for 3' UTR PTGS of different strains of Zika virus: A computational approach. *Nature and Science*, 1, 37–50. doi:10.1017/CBO9781107415324.004
- Millichap, J. G. (2016). Zika virus infection and microcephaly. *Pediatric Neurology Briefs*, 30, 8. doi:10.15844/pedneurbriefs-30-1-7
- Miner, J. J., & Diamond, M. S. (2016). Understanding how zika virus enters and infects neural target cells. *Cell Stem Cell*, 18, 559–560. doi:10.1016/j.stem.2016.04.009
- Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., ... Avšič Županc, T. (2016). Zika virus associated with microcephaly. *New England Journal of Medicine*, 374, 951–958. doi:10.1056/NEJMoa1600651
- Morris, G., & Huey, R. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30, 2785–2791. doi:10.1002/jcc.21256.AutoDock4
- Mumtaz, N., van Kampen, J. J. A., Reusken, C. B. E. M., Boucher, C. A. B., & Koopmans, M. P. G. (2016). Zika virus: Where is the treatment? *Current Treatment Options in Infectious Diseases*, 8, 208–211. doi:10.1007/s40506-016-0083-7
- Munir, S., Saleem, S., Idrees, M., Tariq, A., Butt, S., Rauff, B., ... Awan, Z. (2010). Hepatitis C Treatment: Current and future perspectives. *Virology Journal*, 7, 296. doi:10.1186/1743-422X-7-296
- Nair, P. C., & Miners, J. O. (2014). Molecular dynamics simulations: From structure function relationships to drug discovery. *Silico Pharmacology*, 2(4), 1–4. doi:10.1186/s40203-014-0004-8
- Ostrovskii, V. A., Trifonov, R. E., & Popova, E. A. (2012). Medicinal chemistry of tetrazoles. *Russian Chemical Bulletin*, 61, 768–780. doi:10.1007/s11172-012-0108-4
- Panchaud, A., Stojanov, M., Ammerdorffer, A., & Vouga, M. (2016). Emerging role of Zika virus in adverse fetal and neonatal outcomes. *Clinical Microbiological Reviews*, 29, 659–694. doi:10.1128/CMR.00014-16.Address

- Papageorgiou, L., Loukatou, S., Koumandou, V. L., Makalowski, W., Megalooikonomou, V., Vlachakis, D., & Kossida, S. (2014). Structural models for the design of novel antiviral agents against Greek Goat Encephalitis. *PeerJ*, 2, 1–18. doi:10.7717/peerj.664
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF chimera? A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25, 1605–1612. doi:10.1002/jcc.20084
- Ramharack, P., & Soliman, M. E. S. (2016). Zika virus drug targets: A missing link in drug design and discovery – A route map to fill the gap. *RSC Advances*, 6, 68719–68731. doi:10.1039/C6RA12142J
- Roa, M. (2016). Zika virus outbreak: Reproductive health and rights in Latin America. *The Lancet*, 387, 843. doi:10.1016/S0140-6736(16)00331-7
- Sanner, M. F., Olson, A. J., & Spehner, J. C. (1996). Reduced surface: An efficient way to compute molecular surfaces. *Biopolymers*, 38, 305–320. doi:10.1002/(SICI)1097-0282(199603)38:3<305::AID-BIP4>3.0.CO;2-Y
- Shanmugam, A., Velmurugan, D., & Gromiha, M. M. (2016). Identification of dengue viral RNA-dependent RNA polymerase inhibitor using computational fragment-based approaches and molecular dynamics study. *Journal of Biomolecular Structure & Dynamics*, 34, 1512–1532. doi:10.1080/07391102.2015.1081620
- Shapshak, P., Sinnott, J. T., Somboonwit, C., & Kuhn, J. H. (2016). Zika Virus. In *Global Virology I-Identifying and Investigating Viral Diseases* (Chapter 18, pp. 477–500). New York: Springer Science Business Media. doi:10.1007/978-1-4939-2410-3
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., ... Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539. doi:10.1038/msb.2011.75
- Singh, R. K., Dhama, K., Malik, Y. S., Ramakrishnan, M. A., Karthik, K., Tiwari, R., & Joshi, S. (2016, July). Zika virus – Emergence, evolution, pathology, diagnosis and control: Current global scenario and future perspectives – A comprehensive review. *The Veterinary Quarterly*, 2176, 1–43. doi:10.1080/01652176.2016.1188333
- Soliman, M. E. S. (2013). A hybrid structure/pharmacophore-based virtual screening approach to design potential leads: A computer-aided design of South African HIV-1 subtype C protease inhibitors. *Drug Development Research*, 74, 283–295. doi:10.1002/ddr.21078
- Sweeney, N. L., Hanson, A. M., Mukherjee, S., Ndjomou, J., Geiss, B. J., Steel, J. J., ... Frick, D. N. (2015). Benzothiazole and pyrrolone flavivirus inhibitors targeting the viral helicase. *ACS Infectious Diseases*, 1, 140–148. doi:10.1021/id5000458
- Tambunan, U. S. F., Zahroh, H., Utomo, B. B., & Parikesit, A. A. (2014). Screening of commercial cyclic peptide as inhibitor NS5 methyltransferase of Dengue virus through molecular docking and molecular dynamics simulation. *Bioinformation*, 10, 23–27. doi:10.6026/97320630010023
- Tang, H., Hammack, C., Ogden, S. C., Wen, Z., Qian, X., Li, Y., ... Ming, G. L. (2016). Zika virus infects human cortical neural progenitors and attenuates their growth. *Cell Stem Cell*, 18, 587–590. doi:10.1016/j.stem.2016.02.016
- Troncoso, A. (2016). Zika threatens to become a huge worldwide pandemic. *Asian Pacific Journal of Tropical Biomedicine*, 6, 520–527.
- Trott, O., & Olson, A. J. (2010). AutoDock Vina. *Journal of Computational Chemistry*, 31, 445–461. doi:10.1002/jcc.21334
- Turnell, M. J., Hubert, M. J. P., Maquart, Y. M. V., Guilleu-Guillemette, M., & Leparco-Goff, I. (2016). Late sexual transmission of Zika virus related to. *The Lancet*, 6736, 30775. doi:10.1016/S0140-6736(16)30775-9
- Wang, J., Wang, W., Kollman, P. A., & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, 25, 247–260. doi:10.1016/j.jmgm.2005.12.005
- WHO. (2016). *Zika virus, Microcephaly and Guillain-Barré syndrome* (pp. 1–12). Brazil: Author.
- Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C. C., Pettersen, E. F., ... Ferrin, T. E. (2012). UCSF Chimera, MODELLER, and IMP: An integrated modeling system. *Journal of Structural Biology*, 179, 269–278. doi:10.1016/j.jsb.2011.09.006
- Zmurko, J., Marques, R. E., Schols, D., Verbeken, E., Kaptein, S. J. F., & Neyts, J. (2016). The viral polymerase inhibitor 7-deaza-2-c-methyladenosine is a potent inhibitor of in vitro zika virus replication and delays disease progression in a robust mouse infection model. *PLoS Neglected Tropical Diseases*, 10(5), 1–15. doi:10.1371/journal.pntd.0004695
- Zou, X. W., Liu, Y. C., Hsu, N. S., Huang, C. J., Lyu, S. Y., Chan, H. C., ... Li, T. L. (2014). Structure and mechanism of a nonhaem-iron SAM-dependent C-methyltransferase and its engineering to a hydratase and an O-methyltransferase. *Acta Crystallographica Section D: Biological Crystallography*, 70, 1549–1560. doi:10.1107/S1399004714005239

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

Zika Virus NS5 Protein Potential Inhibitors: An Enhanced
***In Silico* Approach in Drug Discovery**

Pritika Ramharack^a and Mahmoud E. S. Soliman^{a*}

Supplementary Material

Figure S1: The multiple sequence alignment

gi|534286613|pdb|4K6M|A
gi|126030613|pdb|2HFZ|A
gi|985757036|ref|YP_009227205.
gi|985483961|pdb|5CCV|A
gi|453055587|pdb|3VWS|A

DQRGSGQVVITYALNTFTNIAVQLVRLMEAEVIGPQHLEQLPRKNKIAVR
DQRGSGQVVITYALNTFTNLAVQLVRMMEGEGVIGPDDVEKLTGKGPKVR
DQRGSGQVVITYALNTFTNLVQLIRNMEAEVLEMQDLWLLRKPE--KVT
DQRGSGQVGTYGLNTFTNMEAQLIRQMEGEGVLSKADLENP-HPLEKKIT
DQRGSGQVGTYGLNTFTNMEAQLVRQMEGEGVLTADLENP-HLLEKKIT
***** **.******: .**.* **.* *: .: : *

gi|534286613|pdb|4K6M|A
gi|126030613|pdb|2HFZ|A
gi|985757036|ref|YP_009227205.
gi|985483961|pdb|5CCV|A
gi|453055587|pdb|3VWS|A

TWLFENGEERVTRMAISGDDCVVKPLDDRFATALHFLNAMS KVRKDIQEW
TWLSENGEERLSRMAVSGDDCVVKPLDDRFATSLHFLNAMS KVRKDIQEW
RWLQSNQWDRKLRMAVSGDDCVVKPIDDRFAHALRFLNDMGKVRKDTQEW
QWLETGKVERLRMAISGDDCVVKPIDDRFANALLALNDMGKVRKDIQW
QWLETGKVERLRMAISGDDCVVKPIDDRFANALLALNDMGKVRKDIQW
** *: *.**.******:***** :* ** *.****** :*

gi|534286613|pdb|4K6M|A
gi|126030613|pdb|2HFZ|A
gi|985757036|ref|YP_009227205.
gi|985483961|pdb|5CCV|A
gi|453055587|pdb|3VWS|A

KPSHGWHWDWQQVFPFCSNHFQEI VMKDGRSIVVPCRQDELIGRARISPGA
KPSTGWDWQQVFPFCSNHFTELIMKDGRITLVTPCRGQDELIGRARISPGA
KPSTGWSNWEVFPFCSHHFNKLYLKDGRSIVVPCRQDELIGRARISPGA
QPSKGWHWDWQQVFPFCSHHFHELIMKDGRKLVVPCRQDELIGRARISPGA
QPSKGWHWDWQQVFPFCSHHFHELIMKDGRKLVVPCRQDELIGRARISPGA
:*** ** :*:*****:*** : : *****.*** *****:*** **

gi|534286613|pdb|4K6M|A
gi|126030613|pdb|2HFZ|A
gi|985757036|ref|YP_009227205.
gi|985483961|pdb|5CCV|A
gi|453055587|pdb|3VWS|A

GWNVKTACILAKAYAQMWLLLYFHRRDLRLMANAICSAVPVDWVPTGRST
GWNVRDTACILAKSYAQMWLLLYFHRRDLRLMANAICSAVPVNWVPTGRST
GWSIRETACILAKSYAQMWLLLYFHRRDLRLMANAICSAVPVDWVPTGRST
GWSLRETACILGKAYAQMWALMYFHRRDLRLASNAICSAVPVHWVPTSR--
GWSLRETACILGKAYAQMWALMYFHRRDLRLASNAICSAVPVHWVPTSR--
.:*:***.*:***** *:***** :*****.***** *

gi|534286613|pdb|4K6M|A
gi|126030613|pdb|2HFZ|A
gi|985757036|ref|YP_009227205.
gi|985483961|pdb|5CCV|A
gi|453055587|pdb|3VWS|A

WSIHSGKEWMTTEDMLQVWNRVWIEENEWMMDKTPITSWTDVPYVGKRED
WSIHAGGEWMTTEDMLEVWNRVWIEENEWMMDKTPVEKWSVDPYVGKRED
WSIHSGKEWMTTEDMLMVWNRVWIEENDHMDKTPVTKTWDIPYLGKRED
----TTHQWMTTEDMLTVWNRVWIEDNPWMEDKTPVTTWEDVPYLGKRED
WSIHAHQWMTTEDMLTVWNRVWIEENPWMDKTPVTTWENVPYLGKRED
:***** *****:*** * ****: .* :*** *****

gi|534286613|pdb|4K6M|A
gi|126030613|pdb|2HFZ|A
gi|985757036|ref|YP_009227205.
gi|985483961|pdb|5CCV|A
gi|453055587|pdb|3VWS|A

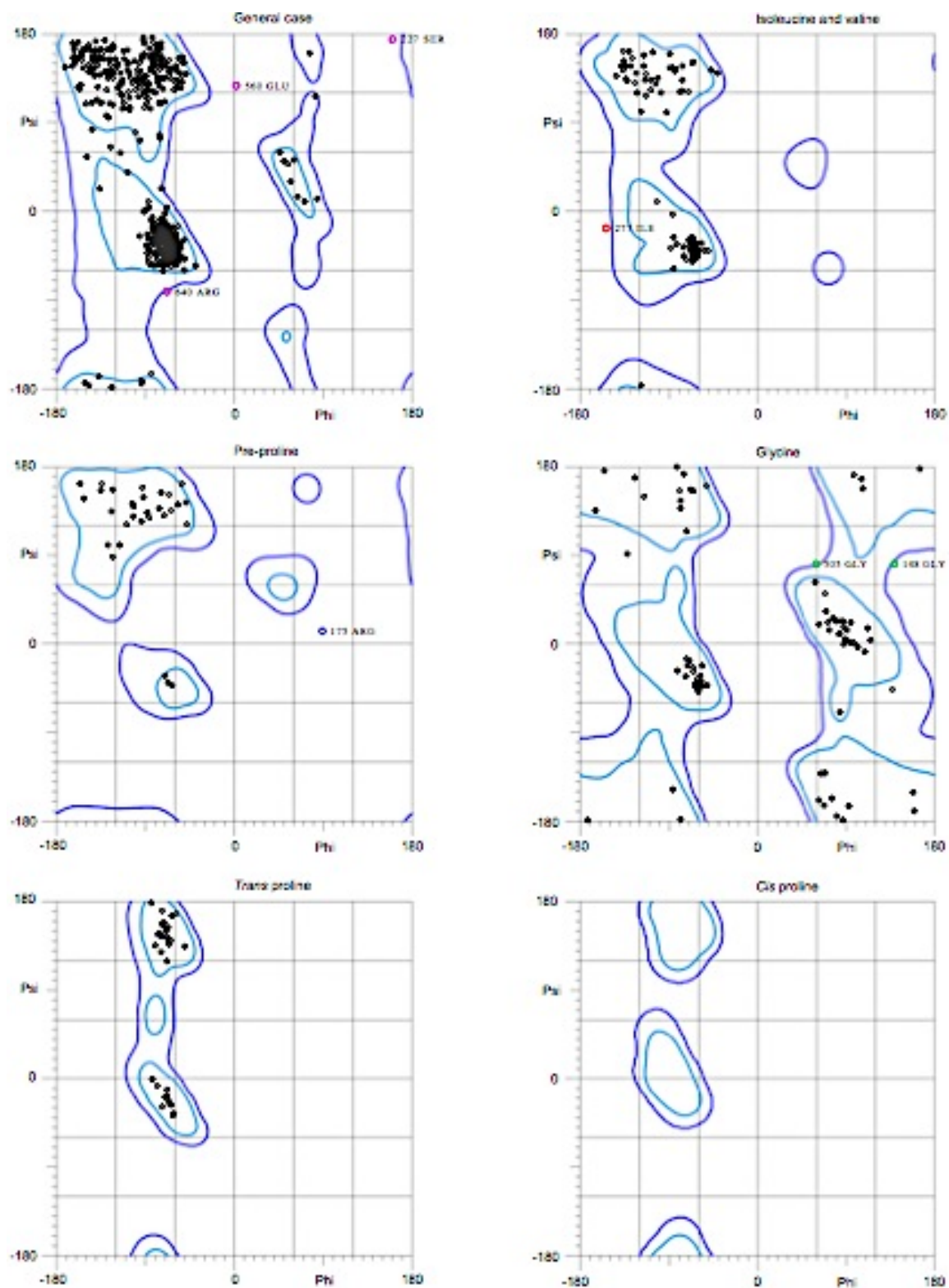
IWCGSLIGTRSRATWAENIYAAINQVRAVIGKE-NYVDYMTSLRRYEDVL
IWCGSLIGTRARATWAENIQVAINQVRSIIGDE-KYVDYMSLKRKYEDTT
LWCGSLIGHRPRTTWAENIKDTVMNVRRIGDEEKYMDYLSQVRYLGEE
QWCGSLIGLTSRATWAQNIITAIQVRSIIGNE-EFLDYMPMSMKRFRKEE
QWCGSLIGLTSRATWAQNIITAIQVRSIIGNE-EFLDYMPMSMKRFRKEE
***** .*:*****:*** ** :***.* :*****.:* *

gi|534286613|pdb|4K6M|A
gi|126030613|pdb|2HFZ|A
gi|985757036|ref|YP_009227205.
gi|985483961|pdb|5CCV|A
gi|453055587|pdb|3VWS|A

IQEDRVIGSSSHHHHHH-
LVEDTVL-----
GSTPGVL-----
ESEGAIWAAALEHHHHHH
ESEGAIW-----

44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67

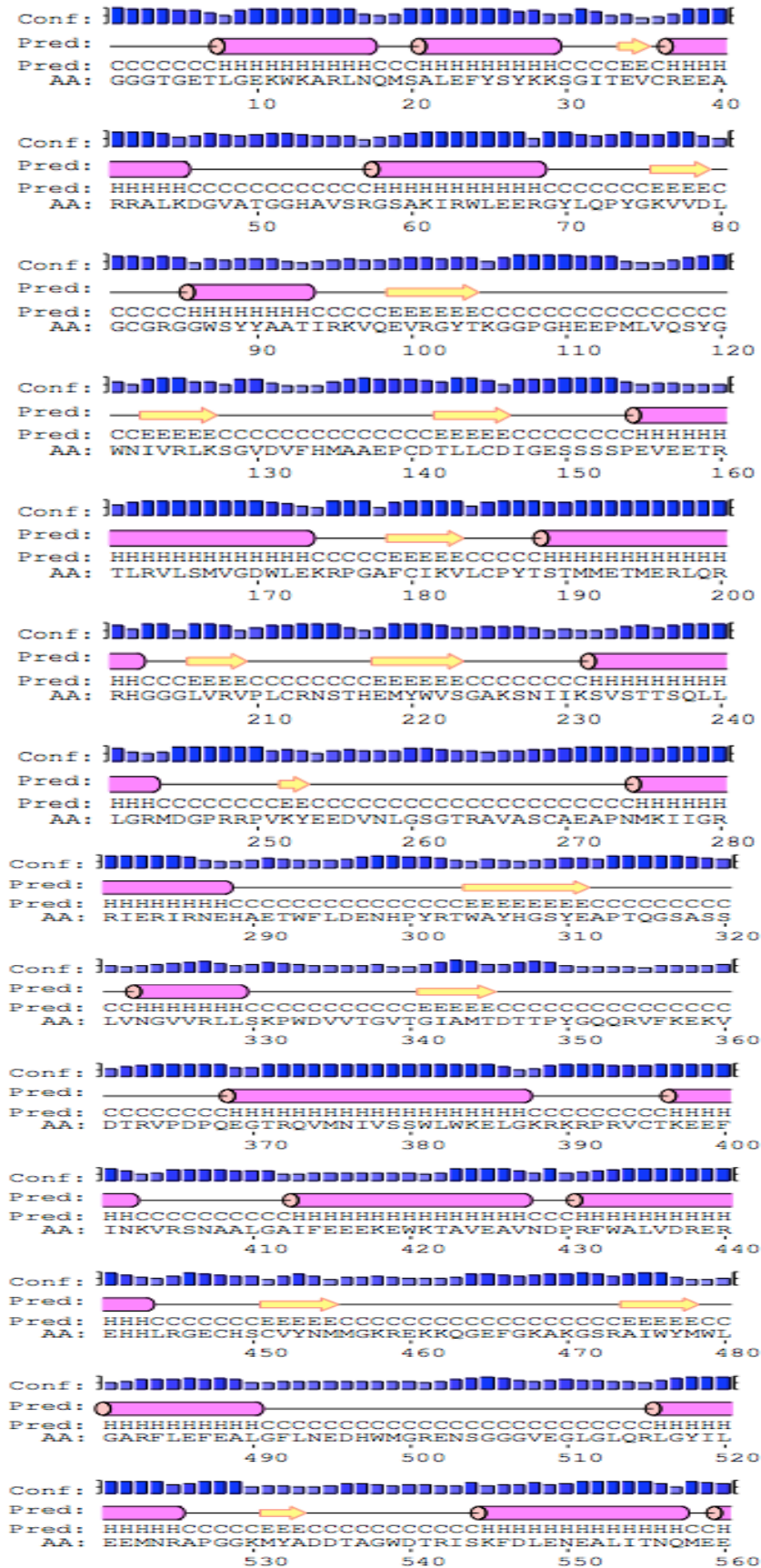
68 **Figure S2:** Ramachandran plot for ZIKV NS5 protein
 69



70
 71
 72
 73

74 97.2% (876/901) of all residues were in favored (98%) regions.
75 99.2% (894/901) of all residues were in allowed (>99.8%) regions.
76
77
78 There were 7 outliers (phi, psi):
79 148 GLY (138.4, 81.3)
80 175 ARG (89.6, 13.4)
81 227 SER (161.0, 175.0)
82 277 ILE (-153.6, -17.7)
83 505 GLY (59.8, 81.0)
84 560 GLU (2.5, 127.7)
85 640 ARG (-68.3, -81.5)
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122

Figure S3: Psipred predicted secondary structure



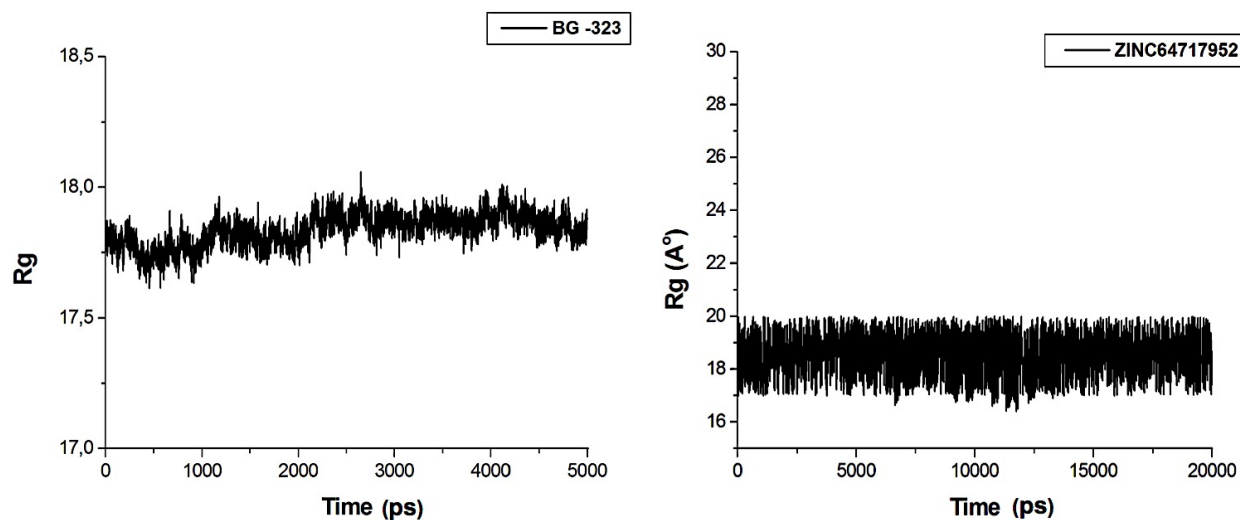


Figure S4.1: Radius of gyration measuring compactness of the MTase protein when in complex with BG323 and ZINC64717952.

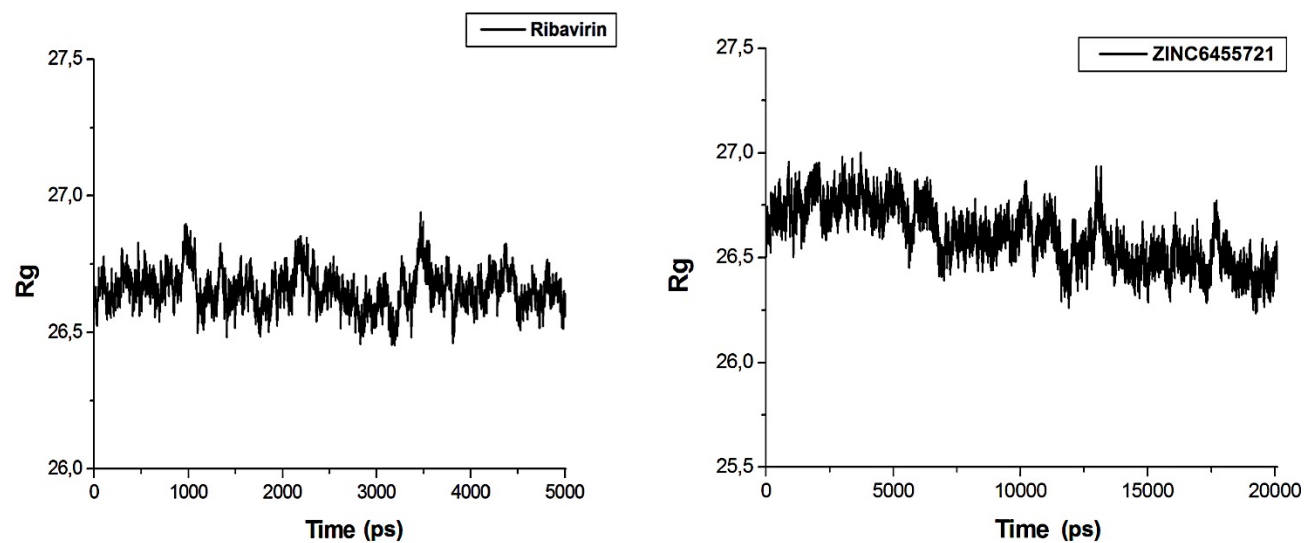



Figure S4.2: Radius of gyration measuring compactness of the RdRp protein when in complex with ribavirin and ZINC6455721.

Cite this: *RSC Adv.*, 2017, 7, 22133

Delving into Zika virus structural dynamics – a closer look at NS3 helicase loop flexibility and its role in drug discovery†

Pritika Ramharack,^a Sofiat Oguntade^a and Mahmoud E. S. Soliman  ^{*abcd}

The Zika virus has emerged as a pathogen of major health concern. The rapid spread of the virus has led to uproar in the medical domain as scientists frantically race to develop effective vaccines and small molecules to inhibit the virus. In the past year, there has been a flood of Zika knowledge published including its characteristics, transmission routes and its role in disease conditions such as microcephaly and Guillain–Barre syndrome. Targeted therapy against specific viral maturation proteins is necessary in halting the replication of the virus in the human host, thus decreasing host–host transmission. This prompted us to investigate the structural properties of the Zika NS3 helicase when bound to ATP-competitive inhibitor, NITD008. In this study, comparative molecular dynamic simulations were employed for APO and bound protein to demonstrate the molecular mechanism of the helicase. Results clearly revealed that NITD008-binding caused significant residue fluctuations at the P-loop compared to the rigid nature of the APO conformation. The NITD008-helicase complex also revealed residues 339–348 to transition from a $_{310}$ -helix to a stable α -helix. These protein fluctuations were verified by investigation of dynamic cross correlation and principal component analysis. The fundamental dynamic analysis presented in this report is crucial in understanding Zika NS3 helicase function, thereby giving insights toward an inhibition mechanism. The information reported on the binding mode at the ATPase active site may also assist in designing effective inhibitors against this detrimental viral target.

Received 2nd February 2017
Accepted 12th April 2017

DOI: 10.1039/c7ra01376k

rsc.li/rsc-advances

1 Introduction

The re-emerging Zika virus (ZIKV) has evolved into a catastrophic epidemic over the past year, with scientific community announcing that the long-term effects associated with the virus will have to be dealt with in the decades to follow.¹ The virus was declared an international public health emergency by the World Health Organization,² based on growing evidence of the virus being linked with congenital neurological diseases such as Guillain–Barre, cranial nerve dysfunction and microcephaly.^{3,4} The ZIKV made its devastating re-appearance in Brazil and has now spread on a global scale, with an estimated 75 countries with reported mosquito-borne ZIKV transmission as of December 2016.⁵

Zika virus is an arthropod-borne *Flavivirus* initially discovered in the Zika forest area of Uganda in 1947.⁶ Of the *Flavivirus* genera, ZIKV is most closely related to the Spondweni virus from the Spondweni group; however, ZIKV shares structural similarities with other *Flaviviruses*, including Dengue virus and West Nile virus.⁷ The ZIKV genome is made up of structural proteins, being the capsid, precursor membrane and envelope form the viral particle and seven non-structural proteins, being NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5, which participate in the replication of the RNA genome, virion assembly and invasion of the innate immune system.^{8–10} In our previous review, we explicated on the key viral target proteins, including the multifunctional viral replication NS3 helicase protein.¹¹ The ZIKV helicase comes from the superfamily helicases, SF2,¹² with the inhibition of either one of the binding sites, the RNA-binding groove or the ATP-binding site (Fig. 1), leading to the virus becoming incapable of sufficient maturation and replication. The structural characteristics of the ZIKV NS3 protein includes three domains: domain I (residues 182–327), domain II (residues 328–480) and domain III (residues 481–617), as well as a P-loop (residues 196–203) which is located at the ATP-binding site of domain I.^{12,13}

The co-crystallization of MnATP^{2−} and RNA with ZIKV helicase, reported by Tian *et al.* (2016) and Cao *et al.* (2016), have paved the way to understanding the mechanism by which these

^aMolecular Modeling and Drug Design Research Group, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa. E-mail: soliman@ukzn.ac.za; Fax: +27 31 260 7872; Tel: +27 31 260 8048

^bSchool of Health Sciences, Pharmaceutical Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

^cDepartment of Pharmaceutical Organic Chemistry, Faculty of Pharmacy, Zagazig University, Zagazig, Egypt

^dCollege of Pharmacy and Pharmaceutical Sciences, Florida Agricultural and Mechanical University, FAMU, Tallahassee, Florida 32307, USA

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c7ra01376k



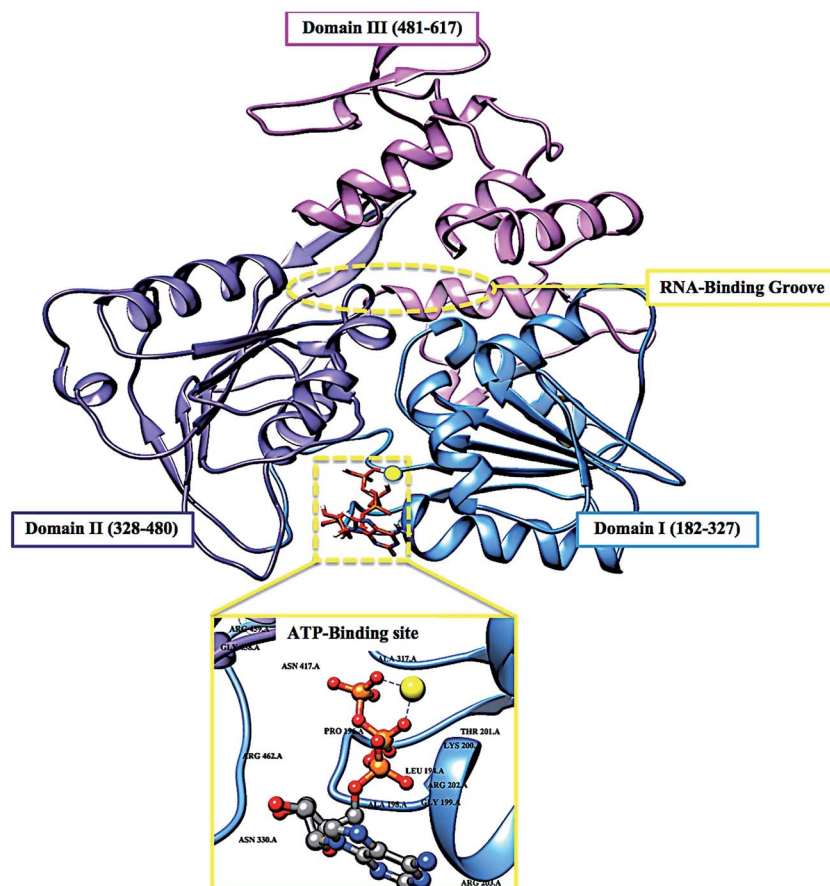


Fig. 1 Cartoon and surface representation of the three domains of the ZIKV helicase and the two active-binding regions (yellow) that form profound hydrophobic cavities in the electrostatic surface area, allowing ATP and ssRNA to bind.

substrates bind to the enzyme, initiating viral RNA replication.^{14,15} Despite the flood of integrated knowledge on ZIKV over the past year, the molecular and structural mechanism for helicase inhibition is yet to be established.¹²

Another battle being fought by researchers is the discovery of new modes of transmission of the virus, from initially being transmitted from vector to host, to now being inclusive of blood transfers from host to host as well as secondary sexual transmission.^{16–18} This has allowed for rapid diffusion of the virus between continents. In the plethora of strategic characteristics of the virus, its ability to target neuronal cells has been one of the most problematic tasks that pharmaceutical chemists have had to overcome.^{19–24} The design of ZIKV inhibitors will not only need to be target-specific, effective and have minimal toxicity, but it will also have to pass through the blood-brain-barrier.²⁵

Although there are currently vaccine clinical trials under way, there are still no FDA approved small molecule inhibitors against the virus.^{26–30} This may be due to a number of reasons including time-consuming experimental testing of large libraries of compounds or minimal literature available on the functionality of the virus in host cells. These possible barriers have prompted us to utilize computational drug design tools, such as molecular dynamic (MD) simulations to explore the conformational landscape of this biological system's ATP-binding region. The crystallographic structures have revealed

evidence of residue mobility, including the rotation of motor domains, however, the precise structural characteristics of the helicase upon small molecule binding, is yet to be determined.^{12,31–36}

In this study we investigate the conformational changes at the ATP-binding region after a 130 ns MD simulation of the free enzyme state as well as a NITD008-bound complex.³⁷ This study will be critical in understanding how the ZIKV NS3 helicase functions structurally, thus aiding in the design of effective, target-specific inhibitors.

2 Computational methods

2.1 System preparation

The ZIKV NS3 helicase in complex with ATP and a magnesium ion (PDB code: 5GJC)¹⁴ was obtained from RSCB Protein Data Bank.³⁸ The 3-D structure of the experimental ZIKV inhibitor, NITD008, was obtained from PubChem³⁹ and prepared on Molegro Molecular Viewer (MMV).⁴⁰ In the ZIKV crystal structure of the ATP-bound helicase, residues A247-S253 were absent, thus the free enzyme (PDB code: 5JMT)¹³ was utilized in the docking of NITD008. Deng *et al.* (2016) reported conclusive *in vivo* evidence of the inhibition of ZIKV by NITD008. The compound is classified as an adenosine nucleoside analog that competitively inhibits ATP, thus sharing an active site.³⁷



2.2 Molecular docking

Molecular docking is a conventional method in computational chemistry which is utilized in the prediction optimized geometric conformations of a ligand within an appropriate binding site.⁴¹ The molecular docking software utilized included Raccoon,⁴² Autodock Graphical user interface supplied by MGL tools⁴³ and AutoDock Vina⁴⁴ with default docking parameters. Prior to docking, Gasteiger charges were added to NITD008 and the non-polar hydrogen atoms were merged to carbon atoms. Water molecules were removed and polar hydrogen was added to the crystal structure of the NS3 helicase. NITD008 was then docked into the ATPase binding pocket of the NS3 helicase (by defining the grid box with spacing of 1 Å and size of 32 × 26 × 30 pointing in x, y and z directions). Due to the lack of experimental data describing ZIKV approved inhibitors, validation of molecular docking based on the lowest energy pose becomes unreliable.⁴⁵ To overcome any experimental bias, the five best conformational poses, based on binding affinities (kcal mol⁻¹), were subjected to MD simulations.

2.3 Molecular dynamic (MD) simulations

Molecular dynamic (MD) simulations provide a robust tool to explore the physical movements of atoms and molecules, thus providing insights on the dynamical evolution of biological systems. The MD simulation was performed using the GPU version of the PMEMD engine provided with the AMBER package, FF14SB variant of the AMBER force field⁴⁶ was used to describe the protein.

ANTECHAMBER was used to generate atomic partial charges for the ligand by utilizing the restrained electrostatic potential (RESP) and the General Amber Force Field (GAFF) procedures. The leap module of AMBER 14 allowed for addition of hydrogen atoms, as well as Na⁺ and Cl⁻ counter ions for neutralization to both the APO- and bound system.

Both systems were then suspended implicitly within an orthorhombic box of TIP3P water molecules such that all atoms were within 10 Å of any box edge.

An initial minimization of 2000 steps was carried out with an applied restraint potential of 500 kcal mol⁻¹ Å⁻² for both solutes, were performed for 1000 steps using a steepest descent method followed by a 1000 steps of conjugate gradients. An additional full minimization of 1000 steps was further carried out by conjugate gradient algorithm without restraint.

A gradual heating MD simulation from 0 K to 300 K was executed for 50 ps, such that the system maintained a fixed number of atoms and fixed volume, *i.e.*, a canonical ensemble (NVT). The solutes within the system are imposed with a potential harmonic restraint of 10 kcal mol⁻¹ Å⁻² and collision frequency of 1.0 ps⁻¹. Following heating, an equilibration estimating 500 ps of the each system was conducted; the operating temperature was kept constant at 300 K. Additional features such as a number of atoms and pressure were also kept constant mimicking an isobaric-isothermal ensemble (NPT). The systems pressure was maintained at 1 bar using the Berendsen barostat.

The total time for the MD simulation conducted was 130 ns. In each simulation the SHAKE algorithm was employed to constrict the bonds of hydrogen atoms. The step size of each simulation was 2 fs and an SPFP precision model was used. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of 1 bar maintained by the Berendsen barostat, a pressure-coupling constant of 2 ps, a temperature of 300 K and Langevin thermostat with collision frequency of 1.0 ps⁻².

2.4 Post-dynamic analysis

The coordinates of the free enzyme and NITD008 complex were each saved every 1 ps and the trajectories were analyzed every 1 ps using PTRAJ, followed by analysis of RMSD, RMSF and radius of gyration using the CPPTRAJ module employed in AMBER 14 suit.

2.4.1 Binding free energy calculations. Binding free energy calculations is an important end point method that may elucidate on the mechanism of binding between a ligand and enzyme, including both enthalpic and entropic contributions.⁴⁷ To estimate the binding affinity of the docked systems, the free binding energy was calculated using the Molecular Mechanics/GB Surface Area method (MM/GBSA).⁴⁸ Binding free energy was averaged over 15 000 snapshots extracted from the 130 ns trajectory. The free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

$$\Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdW}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (4)$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (5)$$

The term E_{gas} denotes the gas-phase energy, which consist of the internal energy E_{int} , Coulomb energy E_{ele} and the van der Waals energies E_{vdW} . The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol} , was estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G . The non-polar solvation energy, G_{SA} , was determined from the solvent accessible surface area (SASA), using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB} , contribution was estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature respectively.

To obtain the contribution of each residue to the total binding free energy profile at the ATPase site, per-residue free energy decomposition was carried out at the atomic level for imperative residues using the MM/GBSA method in AMBER 14 suit.

The system displaying the most favorable binding interaction and energy contributions were subjected to further analysis.



2.4.2 Dynamic cross-correlation analysis (DCC). Dynamic cross correlation is a widespread method in MD simulations in which the correlation coefficients of motions between atoms of a protein may be quantified.⁴⁹ The dynamic cross correlation between the residue-based fluctuations during simulation was calculated using the CPPTRAJ module incorporated in AMBER 14. The formula used to describe dynamic cross correlation is given below:

$$C_{ij} = \frac{\langle \Delta r_i \Delta r_j \rangle}{(\langle \Delta r_i^2 \rangle \langle \Delta r_j^2 \rangle)^{\frac{1}{2}}}$$

The cross-correlation coefficient (C_{ij}) varies within a range of -1 to $+1$ of which the upper and lower limits correspond to a fully correlated and anti-correlated motion during the simulation process. Where, i and j stands for i^{th} and j^{th} residue respectively and Δr_i or Δr_j represents displacement vectors correspond to i^{th} and j^{th} residue respectively. The generated dynamic cross correlation matrix was constructed in Origin software.

2.4.3 Principal component analysis (PCA). Principal component analysis (PCA) is a covariance-matrix-based mathematical technique that is able to demonstrate atomic displacement and the loop dynamics of a protein.⁵⁰ Prior to processing the MD trajectories for PCA, the trajectories of the free enzyme (APO) and the NITD008-bound complex (complex) were stripped of solvent and ions using the PTRAJ module in AMBER 14. The stripped trajectories were then aligned against their corresponding fully minimized structures. PCA was performed for C- α atoms on 900 snapshots each. Using in-house scripts, the first two principal components were calculated and the covariance matrices were generated. The first two principal components (PC1 and PC2) generated from each trajectory were averaged for both the free-enzyme and NITD008-complex. The first two principal components (PC1 and PC2) were computed and a 2×2 covariance matrix were generated using Cartesian coordinates of C α atoms. PC1 and PC2 correspond to first two eigenvectors of covariant matrices. Origin software⁵¹ was used to construct PCA plots.

3 Results and discussion

3.1 NITD008–NS3 helicase complex

3.1.1 Binding of NITD008 with ZIKV helicase. Research into ZIKV inhibitors has been minimal before 2016. However, NITD008, a *Flavivirus* adenosine analogue was evidenced, both

in vitro and *in vivo*, to inhibit ZIKV replication. The adenosine nucleoside analogue competes with natural ATP substrates, which are incorporated into the growing RNA chain. By this substitution, NITD008 is incorporated into the RNA chain, thus terminating the RNA elongation and inhibiting ZIKV maturation.³⁷

Molecular docking has become a major computational tool that is used to predict the orientation of a ligand at a binding site on the receptor. Results from docking often display multiple predicted orientations of the ligand within the active pocket.⁵²

In this study, NITD008 docked at the ATP-binding site in 6 favorable conformations (Fig. S2–S6[†]), with the highest binding-affinity being $-8.2 \text{ kcal mol}^{-1}$. Scoring functions often attempt to reproduce experimental binding affinities, but most software do not always yield the best prediction. Validation of the docked structure with experimentally known drugs was also not possible due to the lack of FDA inhibitors against ZIKV.^{45,53,54}

In an attempt to improve the binding affinity prediction of NITD008, all 6 predicted complexes were subjected to 130 ns molecular dynamic simulations, allowing for more realistic receptor flexibility in an implicit solvent. Each complex was then analyzed using the accurate, MM/GBSA, free binding energy calculation to determine the most favorable pose of NITD008 at the NS3 ATPase active site.^{47,55–57}

3.1.2 Free energy calculations. The total binding free energy for each of the 6 poses of the NITD008–NS3 helicase complex were calculated using the MM/GBSA approach to better understand the various energy contributions within the binding pocket and assess which binding pose would show the most favorable intermolecular interactions at the helicase active site. Per residue decomposition analysis was also assessed and the residue–ligand interaction network of each pose were depicted as “ligplot” maps (Fig. S2–S6[†]). Of the six systems, the pose with the highest docking score, $-8.2 \text{ kcal mol}^{-1}$, showed the most favorable free binding energy ($-55.90 \text{ kcal mol}^{-1}$) supported the molecular docking score, indicating a favorable structural pose of NITD008 at the binding site.

The thermodynamic energy contribution of NITD008 to the total binding free energy of the complex surmounts to the stability of NITD008 in the ATP binding pocket and thus the stability of the complex during the simulation. Table 1 summarizes the free binding energy of the system taking into account the energies of the NS3 helicase and NITD008.

Fig. 2 represents the residue interaction plot of NITD008 within the active site. The active site residues Gly199, Lys200

Table 1 Summary of free binding energy contributions to the NITD008–NS3 helicase system

	Energy components (kcal mol^{-1})				
	ΔE_{vdw}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV helicase	-3429.35 ± 30.09	$-28\ 758.51 \pm 159.37$	$-32\ 187.86 \pm 155.05$	-5121.93 ± 115.09	$-37\ 309.79 \pm 71.27$
NITD008	-4.69 ± 0.85	18.12 ± 5.27	13.43 ± 5.28	-221.12 ± 3.35	-207.68 ± 3.72
Complex	-37.71 ± 4.12	-382.94 ± 28.72	-420.64 ± 28.59	364.75 ± 22.80	-55.90 ± 7.71



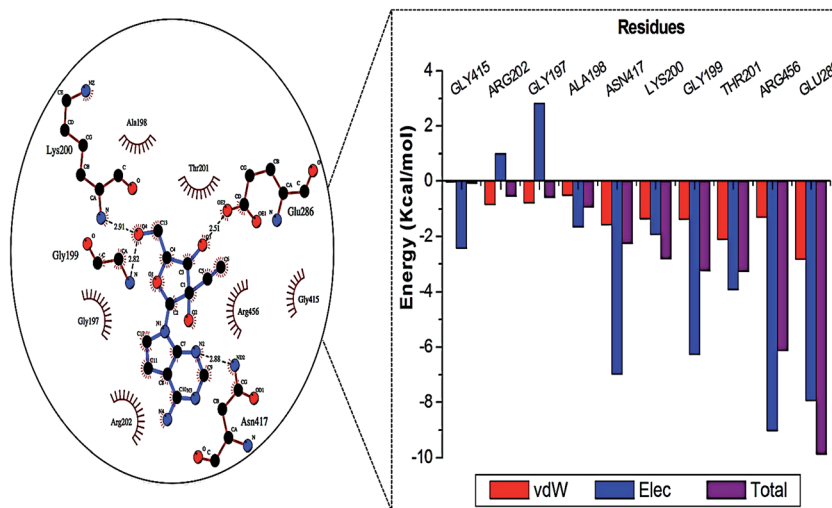


Fig. 2 Energy contributions of the highest interacting residues at the ATPase active site. The residue ligand interaction network illustrates stabilizing hydrophobic interactions pocketing NITD008 at the active site. The highest energy contribution was a hydrogen bond interaction shared between Glu286 and the 3'rd oxygen of the ribose component of NITD008.

and Glu286 formed stable hydrogen bonds with highly electronegative oxygen atoms of NITD008. The residues pocketing NITD008 within the active site included Gly197, Ala198, Gly199, Lys200, Thr201, Arg202, Glu288, Gly415, Asn417 and Arg456.

It was also interesting to note that the most favorable NITD008-pose shared five active residues with the ATP-bound helicase reported by Tian *et al.* (2016). The crystal structure of the ATP-bound helicase showed Lys200 to stabilize the

triphosphate of the ATP.¹⁴ The Lys200 of the NITD008-bound helicase showed a similar stabilizing hydrogen bond with the terminal hydroxyl group located on the ribose of NITD008.

Superimposition of NITD008-docked NS3 helicase with the ATP-NS3 helicase complex demonstrated both compounds to bind in a hydrophilic conformation despite the carbon and acetylene substitutions at *N*-7 of the purine and the 2' position of the ribose, respectively (Fig. 3).

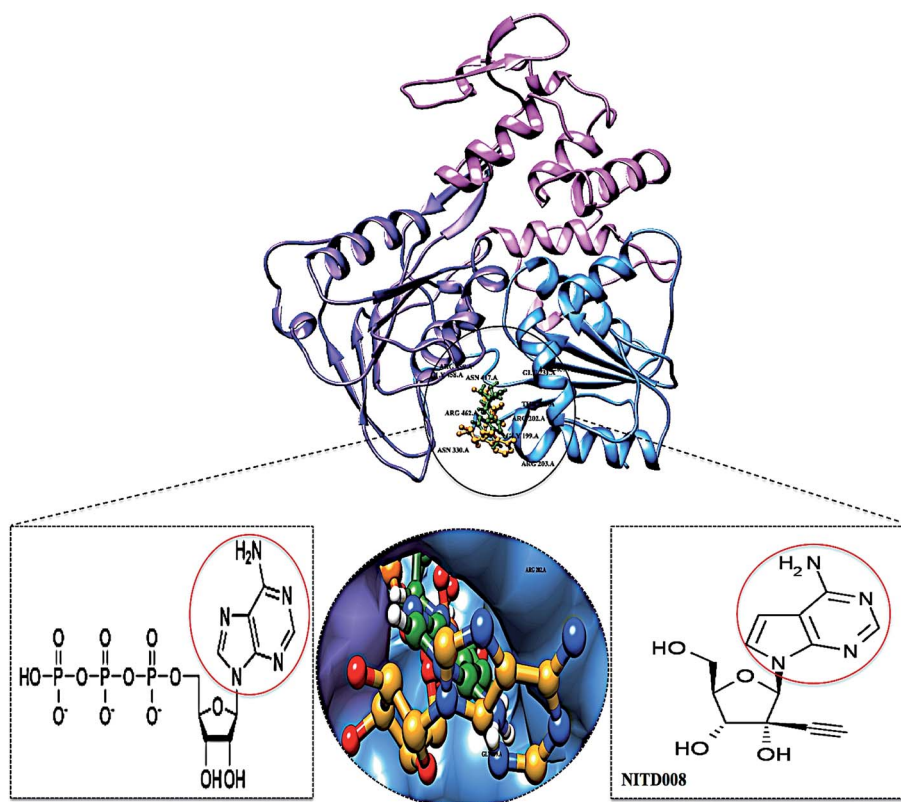


Fig. 3 Superimposed conformation of structurally similar NITD008 and ATP docked at ATPase site of ZIKV NS3 helicase.



The structural similarities between NITD008 and ATP, as well as the active site residue interactions and accurate free-binding energy prompted the further analysis of NITD008-complex.

3.2 Systems stability

The length of a MD simulation is paramount when establishing insights into the structural dynamics of a biological system. With an extended simulation time, a system is able to reach convergence, thus becoming stable. To assure the equilibration of the simulation, the potential energy and temperature were monitored (Fig. S1†). The average potential energy ($-145\,774\text{ kcal mol}^{-1}$) was measured at 300 K, suggesting a stable conformation at this temperature.

3.2.1 Stability of NS3 helicase APO and bound system. The C- α backbone root mean square deviations (RMSD) were monitored throughout the 130 ns MD simulation for both the free (APO) enzyme and the complex. Both systems reached convergence after 60 ns (RMSD deviation $< 2\text{ \AA}$). It can be noted that the C- α backbone atoms in both systems stabilized after a 40 ns time period, although, fluctuations in rigidity did increase during the 47–52 ns period in the NITD008 complex (Fig. 4). This could possibly be due to the occurrence of conformational changes because of the bond interactions taking place between NITD008 and the active site residues as seen in the per-residue energy decomposition.

3.2.2 Conformational fluctuations of the NS3 helicase. To better understand the structural changes that may be occurring upon ligand binding, the root mean square fluctuation (RMSF) of the C- α atoms of each residue in the APO system and NITD008-complex were calculated. Fig. 5 clearly demonstrates greater flexibility of residues of the NITD008-complex when compared to the APO enzyme. Fluctuations take place between

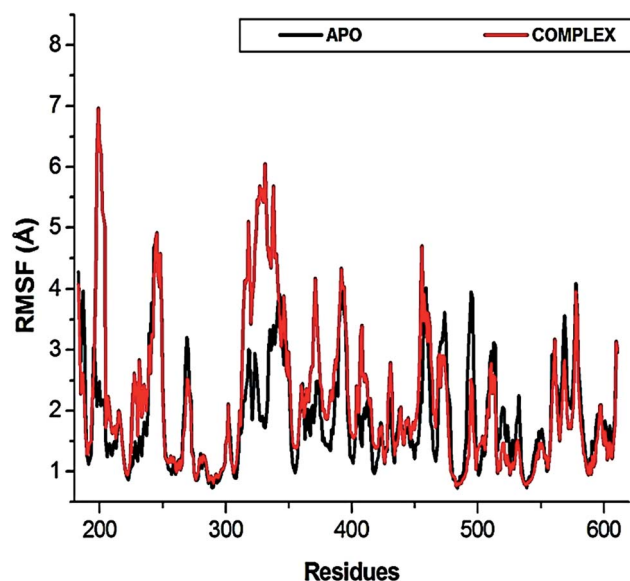


Fig. 5 The RMSF of APO enzyme and NITD008-complex. The structural flexibility in domain I and II is highly attributed to the binding of NITD008 to the ATP-active site. This is substantiated by the average RMSF of the NITD008-complex (2.17 \AA), which is significantly higher than that of the APO enzyme (1.90 \AA).

residues 198–204, which form distinct hydrophobic and hydrogen bond interactions with NITD008 at the active site. This region, the P-loop, is found in all *Flavivirus* helicases and has been shown to have flexibility during binding of ATP.¹⁴ The P-loop adopts structural modifications to accommodate the binding of ATP and Mn^{2+} . This flexibility extends greatly in comparison to the APO enzyme, thus verifying ZIKV P-loop flexibility upon ligand-binding. Other fluctuations occurred in domain II, and I around the ATP-active site, at residues 244–248 and 325–348.

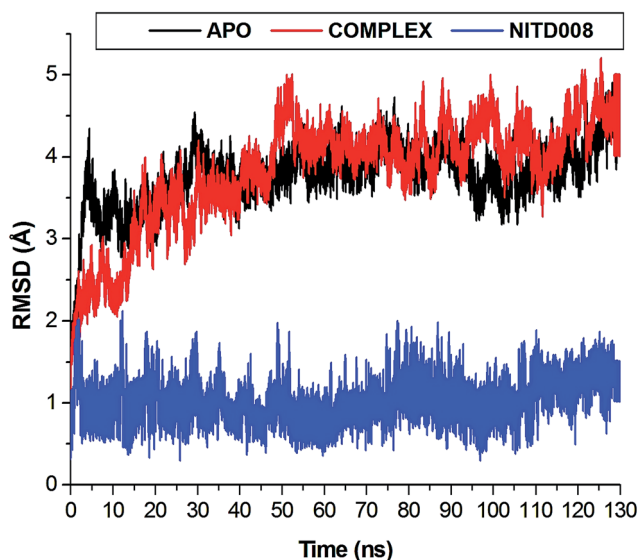


Fig. 4 C- α backbone RMSD for NS3 helicase APO enzyme and NITD-complex conformation. The average C- α RMSD was calculated to be 3.62 \AA and 3.77 \AA , respectively. Increased fluctuations occurred at 47–52 ns in the NITD008-complex.

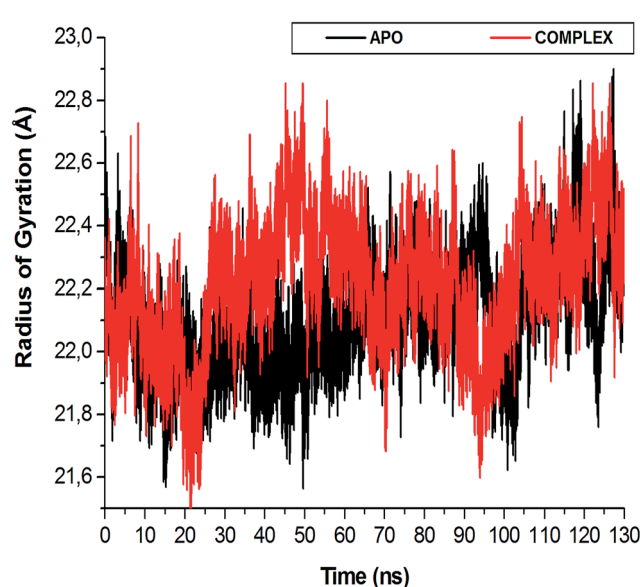


Fig. 6 The radius of gyration (R_g) plot illustrating the difference in enzyme compactness of the NITD008-complex compared to the APO enzyme.



3.2.3 Distribution of atoms around the NS3 helicase backbone. The radius of gyration around the C- α atoms can measure the shape and folding of NS3 helicase before and after NITD008 binding. The radius of gyration measures the distribution of atoms from the center of mass (COM), thus indicating how compact a system is. Both the APO (22.05 Å) and NITD008 (22.17 Å) showed very similar structural compactness, however, there was an atomic distribution in the NITD008-complex from 40–58 ns (Fig. 6). This correlates with the escalated instability of the complex at 47–52 ns demonstrated in the RMSD plot.

The flexibility calculated from the RMSD, RMSF and R_g encouraged us to explore the dynamic structural modifications of the NS3 helicase after NITD008 binding.

3.3 Investigation of the dynamic structural features ATP-active binding region

3.3.1 Loop flexibility and distance metrics. The ZIKV NS3 helicase is made up of three known flexible loops that are common to all *Flaviviruses*: the P-loop (residues 196–203), the RNA-binding loop (residues 244–255) and the β -hairpin loop (residues 431–444). These loops may vary in size depending on the type of virus; however, they all have the same fundamental structural flexibility. The RMSF plot demonstrated major fluctuations at the P-loop as well as the RNA-binding loop, the β -hairpin loop however, showed no significant conformational change compared to the APO enzyme. The plot also illustrated a flexible “325–338” region. Fig. 7 depicts three snapshots of the

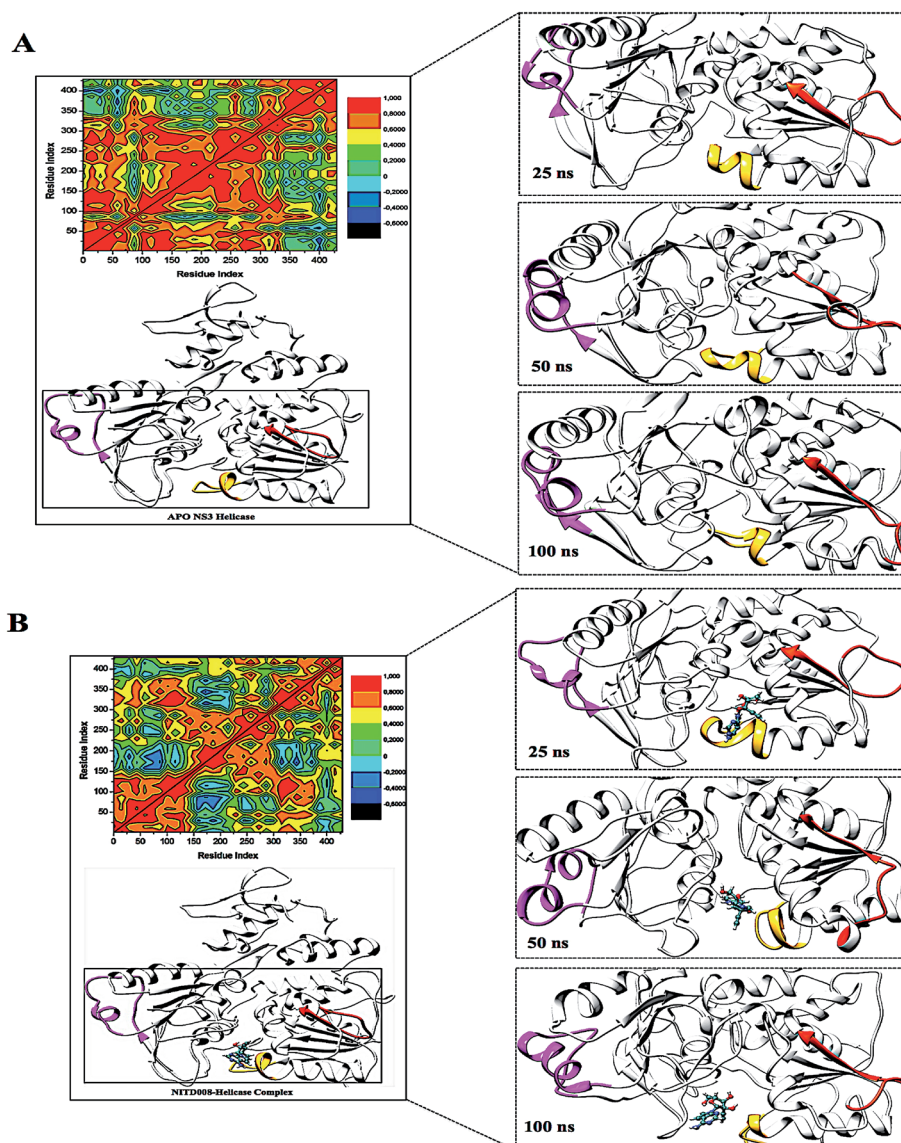


Fig. 7 Structural flexibility of the P-loop (196–203), RNA-binding loop (244–255), and the β 10 helix (339–348) along the trajectory. The RNA-binding loop (orange) showed the loop shifting down in the APO structure but an upward shift in the NITD008-helicase complex. The P-loop (yellow) shifted away from the active site in the bound complex but closed in on the active site when no ligand was present. In the APO structure, the helix-loop-helix stayed, with vibrational movement during the simulation, although, in the bound complex, the β 10 helix (pink) was modified into a α -helix due to ligand motional shifts further into the hydrophobic pocket.



APO enzyme and NITD008-complex, taking at different intervals along the trajectory. Clear conformational shifts are illustrated along the trajectory in both APO and bound systems.

To further investigate the conformational changes of the NS3 helicase upon ligand binding, dynamic cross-correlation matrix (DCCM) analysis was performed at different conformational positions of the C α backbone atoms of the free protein and ligand-bound complex. Highly correlated motions of residues are represented in the red to yellow regions, whereas, the negative/anti-correlated movements of residue C α atoms are represented by blue-navy regions. It is evident from the correlation map that more globally correlated motion is observed in the case of the free protein, confirming conformational shifts after ligand binding. The latter residues of the NS3 helicase,

being residues 500–600, displayed anti-correlated movements in both the APO and bound complex, supporting the residue fluctuations in Fig. 5. Fig. 7 also depicts anti-correlation motions at residues “340–390”, which may be explained by the snapshots, in which, the flexible region in the NITD008-bound complex was converted from a $_310$ -helix to a α -helix.

The P-loop clearly illustrates that when NS3 helicase is in its APO form and exposed to a 130 ns simulation, the P-loop closes on the active site by uncoiling the α -helix at Arg203 to form part of the loop. The loop tip (Ala198) and the adjacent catalytic residue (Gly451) had an average distance of 9.71 Å compared to the NITD008-complex distance of 12.75 Å, whereby, as NITD008 becomes more stable at the active site and forms bond interactions, the P-loop is directed away from NITD008 and a larger

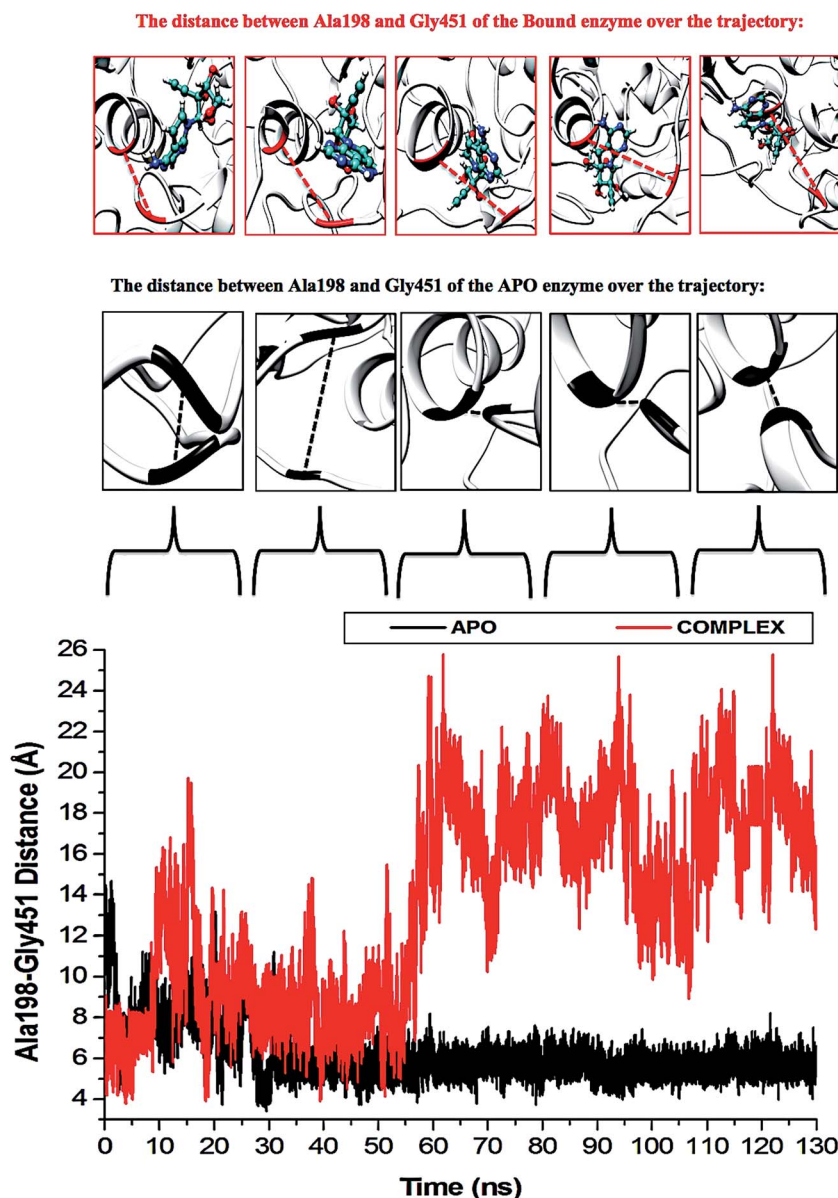


Fig. 8 Residue fluctuations at the P-loop region. The APO enzyme illustrates closing of the loop at the active site due to a vacant hydrophobic pocket. Subsequent to ligand binding and the initiation of stabilizing hydrogen and hydrophobic bond interactions, the P-loop shifts down to accommodate the ligand, thus increasing the size of the hydrophobic pocket.



catalytic space becomes available for the ligand as it forms stable hydrophobic interactions deeper within the hydrophobic pocket (Fig. 8).

The “325–348” region demonstrates opposing conformational modifications between the APO and complex systems compared to that of the P-loop. The distance between the two catalytic residues from the loop tips; residue Ser324 and residue Asn448, measured for the APO and NITD008-complex was 6.34 Å and 8.34 Å, respectively (Fig. 9). The NITD008-complex had a greater distance between the residues due to the unraveling of 2 β -sheets found in domain II. This led to a “325–338” loop shift behind the active site and the “339–348” region being modified from a $_310$ helix to a α -helix (Fig. 7). The $_310$ helix conversion could be due to many reasons including changes in pH, interactions with other proteins and in this case, ligand binding. The

ligand–protein interactions lead to distances between nitrogen and oxygen atoms from the protein backbone to fluctuate and as NITD008 moved further into the hydrophobic pocket, these fluctuations and hydrogen bond conversions caused the $_310$ helix to convert to an α -helix. These changes are important in illustrating the conformational fluctuations upon ligand binding.

3.3.2 Principal component analysis. Conformational transitions of the free protein and NITD008-bound complex were characterized using PCA, a technique that has been widely employed to present experimentally detected conformational variations. Fig. 10 highlights the motional shifts across two principle components in the case of NITD008-bound and unbound NS3 helicase. It is evident that eigenvectors computed from the respective simulations varied immensely between the

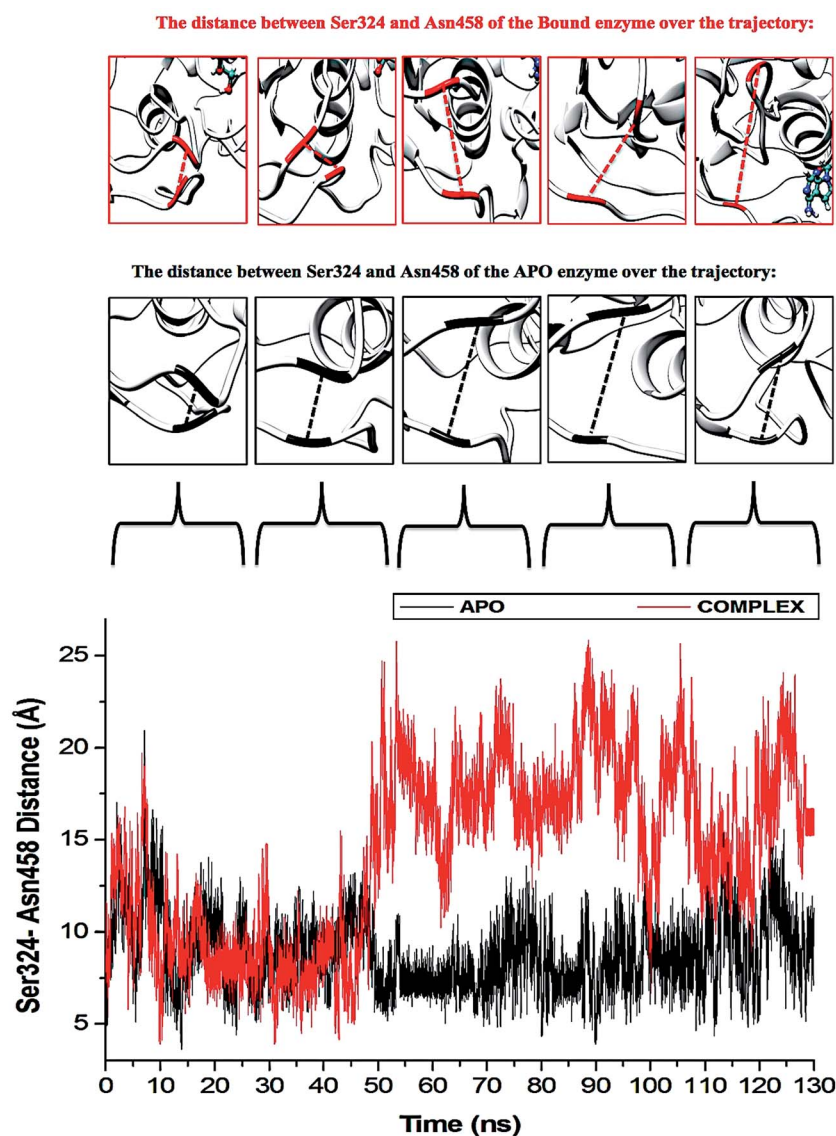


Fig. 9 Residue fluctuations at the “325–348” region. The APO enzyme illustrates widening of the loops of the APO enzyme. The rear loop shifts down as the P-loop closes in on the active site. The largest fluctuation is seen after system stabilization at 40–60 ns. The NITD008-helicase enzyme shows instability in both loops throughout the simulations, although, there was no widening of the loops as the rear loop shifted back rather than downward movement seen in the APO system.



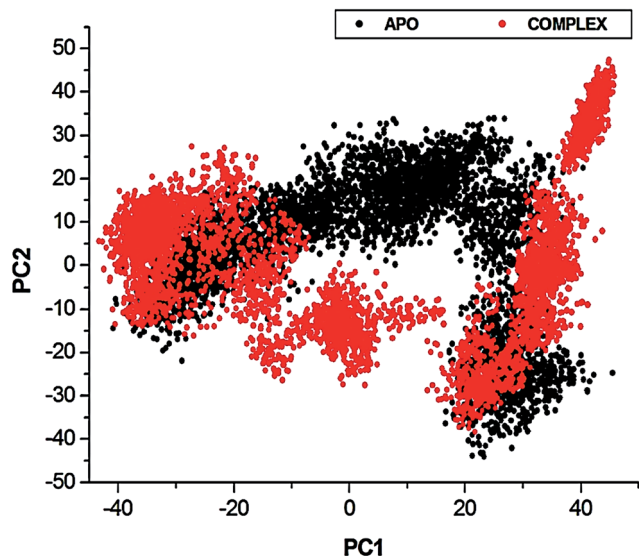


Fig. 10 Projection of Eigen values of the $C\alpha$ backbone, during 130 ns simulation, for APO and NITD008-bound conformations of NS3 helicase along the first two principal components. The X- and Y-axis, PC1 and PC2, respectively, represent a covariance matrix after elimination of eigenvectors (rotational movements). Each point between the single-directional motions represents a unique conformation during the simulation, whereby, similar structural conformations overlap in the graph.

two systems, further elaborating on the dynamic conformational fluctuations from free to ligand-bound protein. The unbound system shows restricted structural motions of residue $C\alpha$ atoms, whereby the NITD008-bound system shows a larger spatial occupancy, thus substantiating the rigidity of the systems, illustrating greater distribution of the atoms around the center of mass and the system stability deviations for the NITD008-bound system. Correlation from analysis of both the free and bound protein demonstrates structural loop flexibility after binding of NITD008 to the ATPase active site.

4 Conclusion

The detailed MD analyses provided in this report demonstrate the structural alterations in ZIKV NS3 helicase loop flexibility subsequent to binding of potent inhibitor, NITD008.³⁷ Molecular simulations revealed profound motional shifts of the ZIKV P-loop at the ATPase active site. This flexibility was revealed in the RMSF analysis and verified by graphical investigation of the loop at different time intervals during the simulation. Investigation into the dynamic cross-correlation of the unbound and bound systems as well as a plot of conformational poses along the first two principal components resulted in strongly significant structural flexibility of the NITD008–NS3 helicase system compared to the rigid unbound protein. The P-loop has demonstrates similar motional shifts in other *Flaviviruses* as well as in ZIKV, when natural substrate, ATP binds at the active site. The competitive inhibitor, NITD008, has been proven to effectively constrain ZIKV replication both *in vitro* and *in vivo*.

Complex stability measured through the 130 ns simulation showed consistency of NITD008 at the ATPase active site and binding free energy calculations and residue–ligand networks revealed strong stabilizing hydrophobic and hydrogen bond interactions pocketing NITD008 in the active site. Further conformational changes were illustrated by the “325–338” loop shift behind the active site and the “339–348” region being modified from a fluctuating $_{310}$ helix to a more stable α -helix.

Crystallographic studies have identified the P-loop, specifically Lys200, to be critical in stabilizing the triphosphate moiety of an NTP, thus allowing flexibility upon ligand binding and activation.^{12–14} To augment these key findings, Lys200 showed strong hydrogen bonds with the NTP-analogue, NITD008. Other active-hotspot residues included P-loop residues: Gly197–Arg202, Ala198, Glu286, Gly415, Asn417 and Arg456. The insights demonstrating the above binding landscape of the ZIKV NS3 helicase will aid researchers in the identification of targeted-small molecule inhibitors through structure based drug design and to utilize pharmacophore models in screening for effective drugs with minimal toxicity.

Future experimental analysis is needed to fully understand these loop shifts toward inhibition of the enzyme as well as investigations into possible mutational resistance as seen in other *Flavivirus* helicase NTPase sites.

Acknowledgements

The authors acknowledge the National Research Foundation for their financial support (UID: 102103) and the Center for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

Notes and references

- 1 L. Broxmeyer and R. Kanjhan, *Mod. Res. Inflammation*, 2016, 5, 20–30.
- 2 Centers for Disease Control, Centers Dis. Control Prev. Zika Virus Home, 2016, 1–12.
- 3 S. A. Rasmussen, D. J. Jamieson, M. A. Honein and L. R. Petersen, *N. Engl. J. Med.*, 2016, 374, 1981–1987.
- 4 A. M. Palomo, *J. Publ. Health. Pol.*, 2016, 37, 133–135.
- 5 WHO, *Zika Situation Report*, 2016, pp. 1–12.
- 6 O. Faye, C. C. M. Freire, A. Iamarino, O. Faye, J. V. C. de Oliveira, M. Diallo, P. M. Zanotto and A. A. Sall, *PLoS Neglected Trop. Dis.*, 2014, 8, 1–10.
- 7 R. Tilak, S. Ray, V. W. Tilak and S. Mukherji, *Medical Journal Armed Forces India*, 2016, 72, 157–163.
- 8 C. G. Noble, Y. L. Chen, H. Dong, F. Gu, S. P. Lim, W. Schul, Q. Y. Wang and P. Y. Shi, *Antiviral Res.*, 2010, 85, 450–462.
- 9 M. Mahfuz, A. Khan, H. Al Mahmud, M. Hasan, A. Parvin, N. Rahman and S. M. B. Rahman, *Indian J. Pharm. Biol. Res.*, 2014, 2, 44–57.
- 10 C. Zanluca, C. N. Duarte and D. Santos, *Microbes Infect.*, 2016, 18, 295–301.
- 11 P. Ramharack and M. E. S. Soliman, *RSC Adv.*, 2016, 6, 68719–68731.



- 12 R. Jain, J. Coloma, A. Garcia-Sastre and A. K. Aggarwal, *Nat. Struct. Mol. Biol.*, 2016, **2**, 1–4.
- 13 H. Tian, X. Ji, X. Yang, W. Xie, K. Yang, C. Chen, C. Wu, H. Chi, Z. Mu, Z. Wang and H. Yang, *Protein Cell*, 2016, **7**, 450–454.
- 14 H. Tian, X. Ji, X. Yang, Z. Zhang, Z. Lu, K. Yang, C. Chen, Q. Zhao, H. Chi, Z. Mu, W. Xie, Z. Wang, H. Lou, H. Yang and Z. Rao, *Protein Cell*, 2016, **7**, 562–570.
- 15 X. Cao, Y. Li, X. Jin, Y. Li, F. Guo and T. Jin, *Nucleic Acids Res.*, 2016, **44**, 10505–10514.
- 16 E. D'Ortenzio, S. Matheron, X. de Lamballerie, B. Hubert, G. Piorkowski, M. Maquart, D. Descamps, F. Damond, Y. Yazdanpanah and I. Leparac-Goffart, *N. Engl. J. Med.*, 2016, **374**, 2195–2198.
- 17 A. C. Gourinat, O. O. Connor, E. Calvez, C. Goarant and M. Dupont-Rouzeyrol, *Emerging Infect. Dis.*, 2015, **21**, 84–86.
- 18 M. J. Turmel, M. J. P. Hubert, Y. M. V. Maquart, M. Guillou-Guillemette and I. Leparac-Goff, *Lancet*, 2016, **6736**, 2501.
- 19 A. R. Plourde and E. M. Bloch, *Emerging Infect. Dis.*, 2016, **22**, 1–15.
- 20 J.-M. Anaya, C. Ramirez-Santana, I. Salgado-Castaneda, C. Chang, A. Ansari, M. E. Gershwin, R. Martinez, J. Bhatnagar, M. Keating, L. Silva-Flannery, A. Muehlenbachs, J. Gary, C. Woods, A. Parker, B. Wakerley, A. Uncini, N. Yuki, J. Anaya, Y. Shoenfeld, A. Rojas-Villarraga, R. Levy, M. Dalakas, B. Wakerley, N. Yuki, S. Kivity, M. Arango, M. Ehrenfeld, O. Tehori, Y. Shoenfeld, J. Anaya, A. Denman, B. Rager-Zisman, T. Kolter, H. Willison, N. Yuki, R. Lardone, N. Yuki, F. Irazoqui, G. Nore, I. Kostovic, R. Ghiulai, M. Sarbu, Z. Vukelic, C. Ilie, A. Zamfir, T. Bell, E. Field, H. Narang, O. Faye, C. Freire, A. Iamarino, O. Faye, J. Oliveira, M. Diallo, A. Nahmias, S. Nahmias and D. Danielsson, *BMC Med.*, 2016, **14**, 1–3.
- 21 N. L. Bayless, R. S. Greenberg, T. Swigut, J. Wysocka and C. A. Blish, *Cell Host Microbe*, 2016, **20**, 423–428.
- 22 D. Olganier, M. Muscolini, C. B. Coyne, M. S. Diamond and J. Hiscott, *DNA Cell Biol.*, 2016, **35**, 367–372.
- 23 J. B. Brault, C. Khou, J. Basset, L. Coquand, V. Fraiser, M. P. Frenkiel, B. Goud, J. C. Manuguerra, N. Pardigon and A. D. Baffet, *EBioMedicine*, 2016, **10**, 71–76.
- 24 H. Li, L. Saucedo-Cuevas, J. A. Regla-Nava, G. Chai, N. Sheets, W. Tang, A. V. Terskikh, S. Shresta and J. G. Gleeson, *Cell Stem Cell*, 2016, **19**, 593–598.
- 25 T. J. Nowakowski, A. A. Pollen, E. Di Lullo, C. Sandoval-Espinosa, M. Bershteyn and A. R. Kriegstein, *Cell Stem Cell*, 2016, **18**, 591–596.
- 26 J. Cohen, *Science*, 2016, **351**, 543–544.
- 27 E. Kim, G. Erdos, S. Huang, T. Kenniston, L. D. Falo and A. Gambotto, *EBioMedicine*, 2016, **13**, 315–320.
- 28 T. C. Pierson and B. S. Graham, *Cell*, 2016, **167**, 625–631.
- 29 G. W. A. Dick, S. F. Klitchen and A. J. Haddow, *Trans. R. Soc. Trop. Med. Hyg.*, 1969, **63**, 708–737.
- 30 R. W. Malone, J. Homan, M. V. Callahan, J. Glasspool-Malone, L. Damodaran, A. D. B. Schneider, R. Zimler, J. Talton, R. R. Cobb, I. Ruzic, J. Smith-Gagen, D. Janies, J. Wilson, D. Hone, S. Hone, S. Bavari, V. Soloveva and S. Weaver, *PLoS Neglected Trop. Dis.*, 2016, **10**, 1–26.
- 31 A. N. Hazin, A. Poretti, D. Di Cavalcanti Souza Cruz, M. Tenorio, A. van der Linden, L. J. Pena, C. Brito, L. H. V. Gil, D. de Barros Miranda-Filho, E. T. d. A. Marques, C. M. Turchi Martelli, J. G. B. Alves and T. A. Huisman, *N. Engl. J. Med.*, 2016, **374**, 2193–2195.
- 32 N. Gruba, J. I. Rodriguez Martinez, R. Grzywa, M. Wysocka, M. Skorenski, M. Burmistrz, M. Lecka, A. Lesner, M. Sienczyk and K. Pyrc, *FEBS Lett.*, 2016, **590**, 3459–3468.
- 33 B. D. Cox, R. A. Stanton and R. F. Schinazi, *Antiviral Chem. Chemother.*, 2016, **24**, 118–126.
- 34 J. Lei, G. Hansen, C. Nitsche, C. D. Klein, L. Zhang and R. Hilgenfeld, *Science*, 2016, **353**, 503–505.
- 35 D. Luo, T. Xu, R. P. Watson, D. Scherer-Becker, A. Sampath, W. Jahnke, S. S. Yeong, C. H. Wang, S. P. Lim, A. Strongin, S. G. Vasudevan and J. Lescar, *EMBO J.*, 2008, **27**, 32090–33219.
- 36 A. V. Chernov, S. A. Shiryayev, A. E. Aleshin, B. I. Ratnikov, J. W. Smith, R. C. Liddington and A. Y. Strongin, *J. Biol. Chem.*, 2008, **283**, 17270–17278.
- 37 Y. Q. Deng, N. N. Zhang, C. F. Li, M. Tian, J. N. Hao, X. P. Xie, P. Y. Shi and C. F. Qin, *Open Forum Infect. Dis.*, 2016, **3**, 1–4.
- 38 H. M. Berman, T. Battistuz, T. N. Bhat, W. F. Bluhm, E. Philip, K. Burkhardt, Z. Feng, G. L. Gilliland, L. Iype, S. Jain, P. Fagan, J. Marvin, D. Padilla, V. Ravichandran, N. Thanki, H. Weissig and J. D. Westbrook, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2002, **58**, 899–907.
- 39 S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang and S. H. Bryant, *Nucleic Acids Res.*, 2016, **44**, 1202–1213.
- 40 S. Kusumaningrum, E. Budianto, S. Kosela, W. Sumaryono and F. Juniarti, *J. Appl. Pharm. Sci.*, 2014, **4**, 47–53.
- 41 Z. Yang, K. Lasker, D. Schneidman-Duhovny, B. Webb, C. C. Huang, E. F. Pettersen, T. D. Goddard, E. C. Meng, A. Sali and T. E. Ferrin, *J. Struct. Biol.*, 2012, **179**, 269–278.
- 42 S. Cosconati, S. Forli, A. L. Perryman, R. Harris, D. S. Goodsell and A. J. Olson, *Expert Opin. Drug Discovery*, 2010, **5**, 597–607.
- 43 M. F. Sanner, *Autodock4 and AutoDock-Tools4: automated docking with selective receptor flexibility*, Scripps Res. Inst., 2008, vol. 26, pp. 1–12.
- 44 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 445–461.
- 45 N. Huang, B. K. Shoichet and J. J. Irwin, *J. Med. Chem.*, 2012, **49**, 6789–6801.
- 46 P. C. Nair and J. O. Miners, *In Silico Pharmacology*, 2014, **2**, 1–4.
- 47 M. Ylilauri and O. T. Pentikäinen, *J. Chem. Inf. Model.*, 2013, **53**, 2626–2633.
- 48 T. Hou, J. Wang, Y. Li and W. Wang, *J. Chem. Inf. Model.*, 2011, **51**, 69–82.
- 49 V. Gosu and S. Choi, *Sci. Rep.*, 2014, **4**, 1–13.
- 50 A. M. Martinez and A. C. Kak, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2001, **23**, 228–233.
- 51 E. Seifert, *J. Chem. Inf. Model.*, 2014, **54**, 1552.
- 52 D. Ramirez and J. Caballero, *Int. J. Mol. Sci.*, 2016, **17**, 1–15.



- 53 X.-Y. Meng, H.-X. Zhang, M. Mezei and M. Cui, *Curr. Comput.-Aided Drug Des.*, 2011, **7**, 146–157.
- 54 L. G. Ferreira, R. N. Dos Santos, G. Oliva and A. D. Andricopulo, *Molecular docking and structure-based drug design strategies*, 2015, vol. 20.
- 55 P. A. Greenidge, C. Kramer, J. C. Mozziconacci and R. M. Wolf, *J. Chem. Inf. Model.*, 2013, **53**, 201–209.
- 56 J. M. Hayes and G. Archontis, *Zika Virus outside Africa*, InTech, 2011, pp. 171–190.
- 57 F. Godschalk, S. Genheden, P. Söderhjelm and U. Ryde, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7731–7739.



Delving into Zika Virus Structural Dynamics- A Closer look at NS3 Helicase Loop flexibility and its Role in Drug Discovery

Pritika Ramharack^A, Sofiat Oguntade^A Mahmoud E. S. Soliman^{A*}

^AMolecular Modeling and Drug Design Research Group, School of Health Sciences,
University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

Supplementary Data

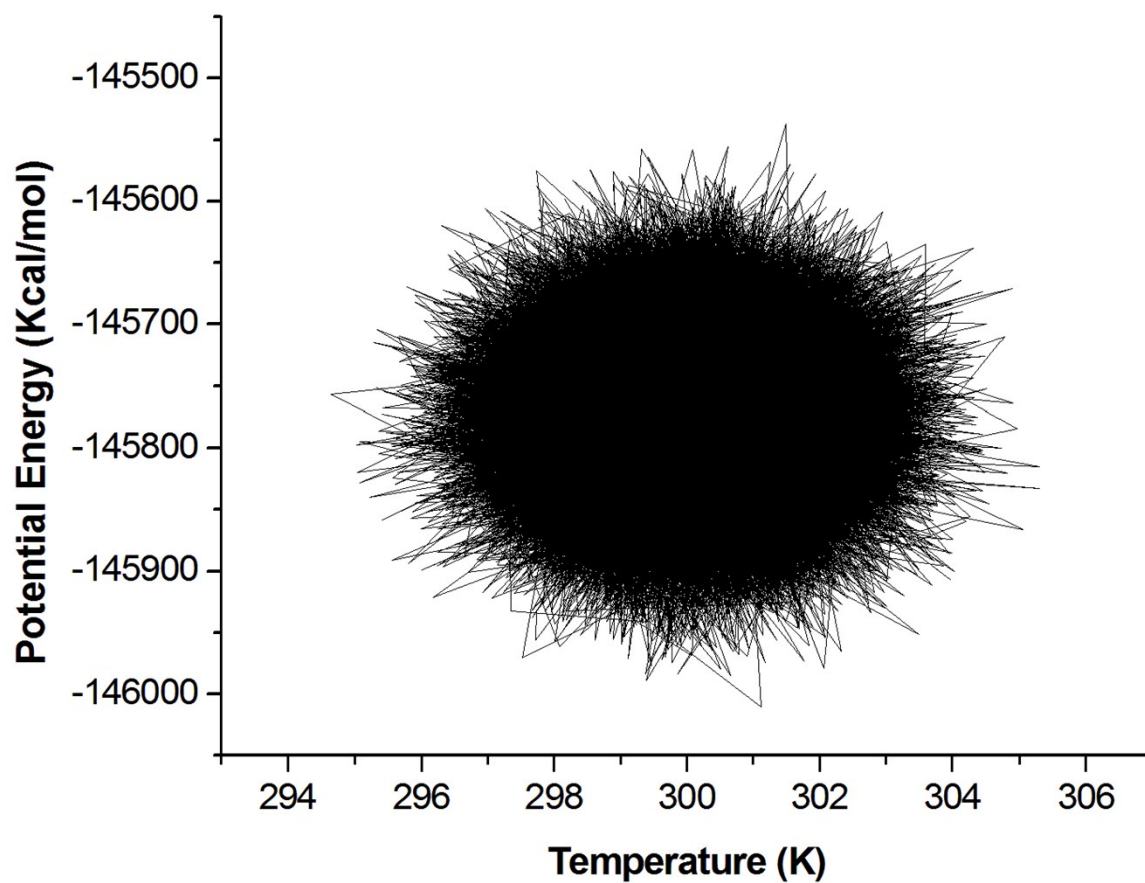


Figure S1: Potential Energy Fluctuations of the NITD008-NS3 Helicase System at varying temperatures during the 100ns simulation. The average temperature of the system was 300K and the average potential energy was -145774 kcal/mol.

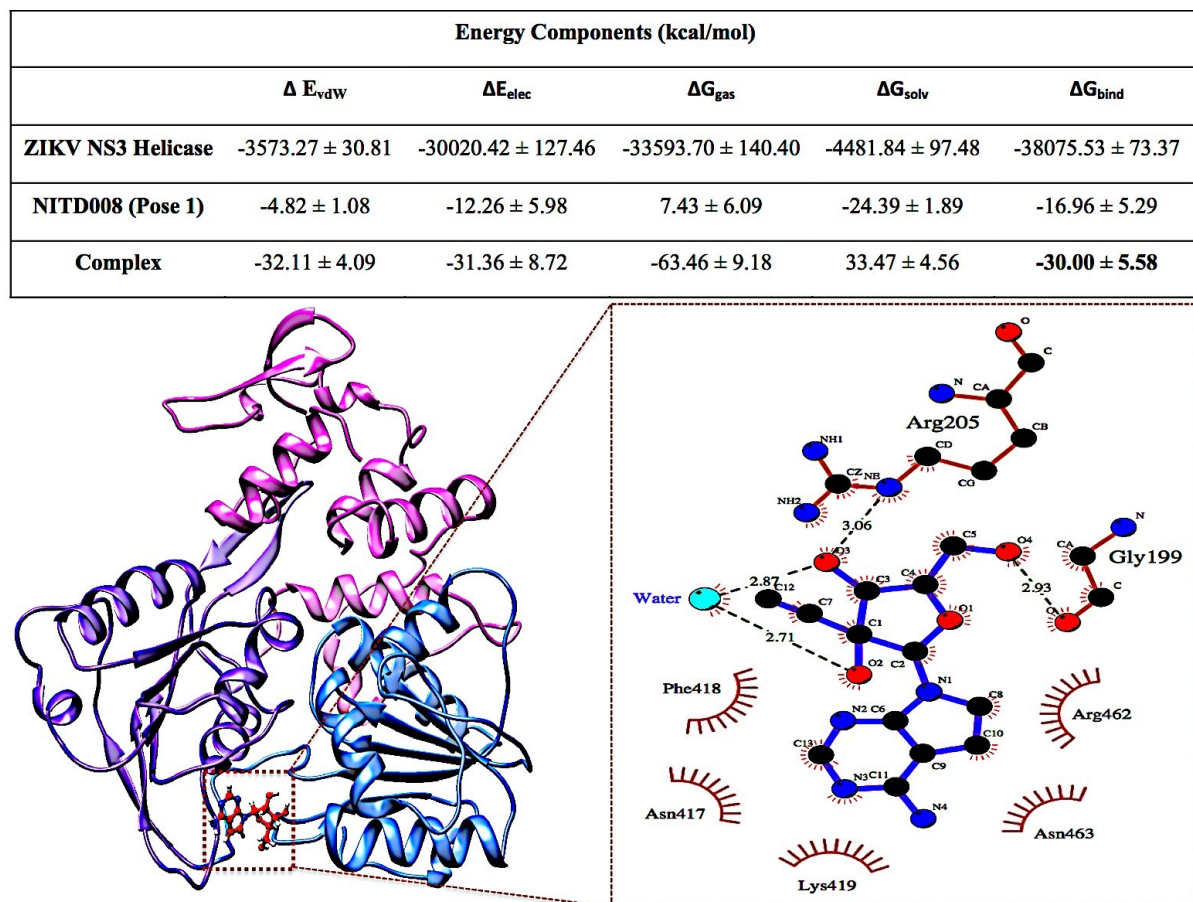


Figure S2: Complex of NITD008-NS3 Helicase with a Docking score of -7.7 kcal/mol. MM/GBSA calculations yielded a result of -30.00 kcal/mol. The ligand shifted further out of the hydrophobic pocket after 150ns of the simulation. This may possibly be due to the ligand not interacting with the stabilizing residues of the P-loop.

Energy Components (kcal/mol)					
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV NS3 Helicase	-3567.30 ± 31.08	-29710.32 ± 175.80	-33277.62 ± 178.68	-4790.36 ± 137.78	-38067.98 ± 76.12
NITD008 (Pose 2)	-5.47 ± 1.01	-6.34 ± 7.85	0.87 ± 7.35	-27.84 ± 3.89	-26.97 ± 5.17
Complex	-24.12 ± 2.75	-9.35 ± 7.29	-33.47 ± 6.78	19.80 ± 5.62	-13.67 ± 3.01

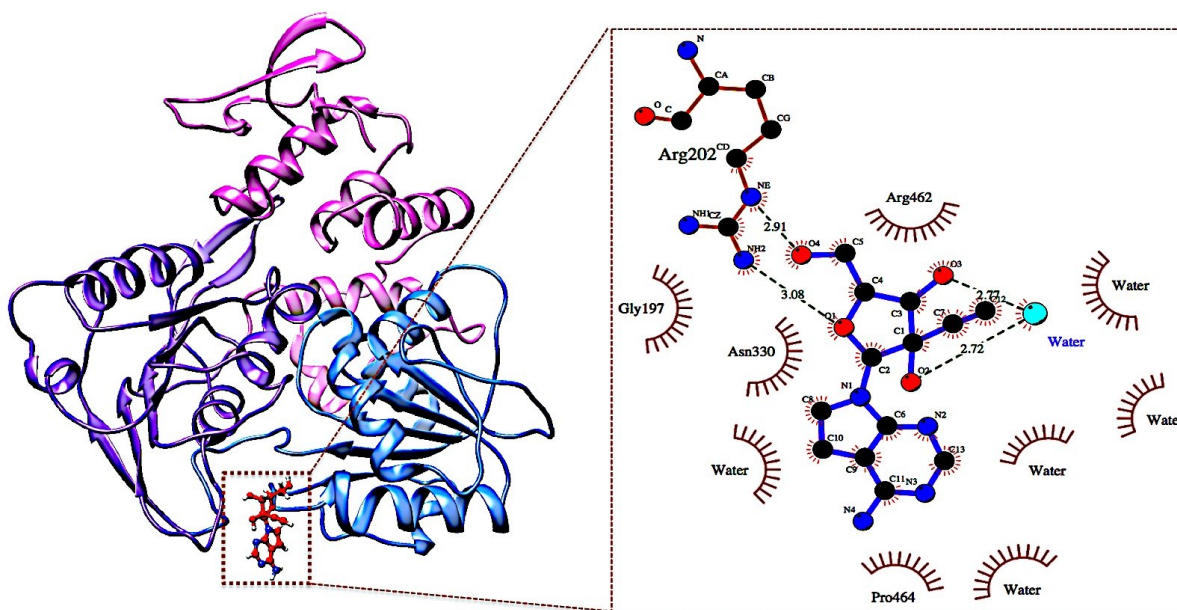


Figure S3: Complex of NITD008-NS3 Helicase with a Docking score of -7.6 kcal/mol. MM/GBSA calculations yielded a result of -13.67 kcal/mol. The ligand docked out of the hydrophobic pocket and during the simulation, due to the lack of stabilizing interactions, the ligand moved further out of the active site and into the solvent.

Energy Components (kcal/mol)					
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV NS3 Helicase	-3564.79 ± 26.89	-29878.89 ± 131.37	-33443.68 ± 133.78	-4633.21 ± 88.03	-38076.89 ± 68.96
NITD008 (Pose 3)	-5.56 ± 0.60	10.53 ± 6.58	4.96 ± 6.51	-28.45 ± 3.51	-23.49 ± 4.78
Complex	-17.75 ± 4.65	-14.45 ± 7.46	-32.21 ± 8.76	20.35 ± 5.95	-11.86 ± 6.37

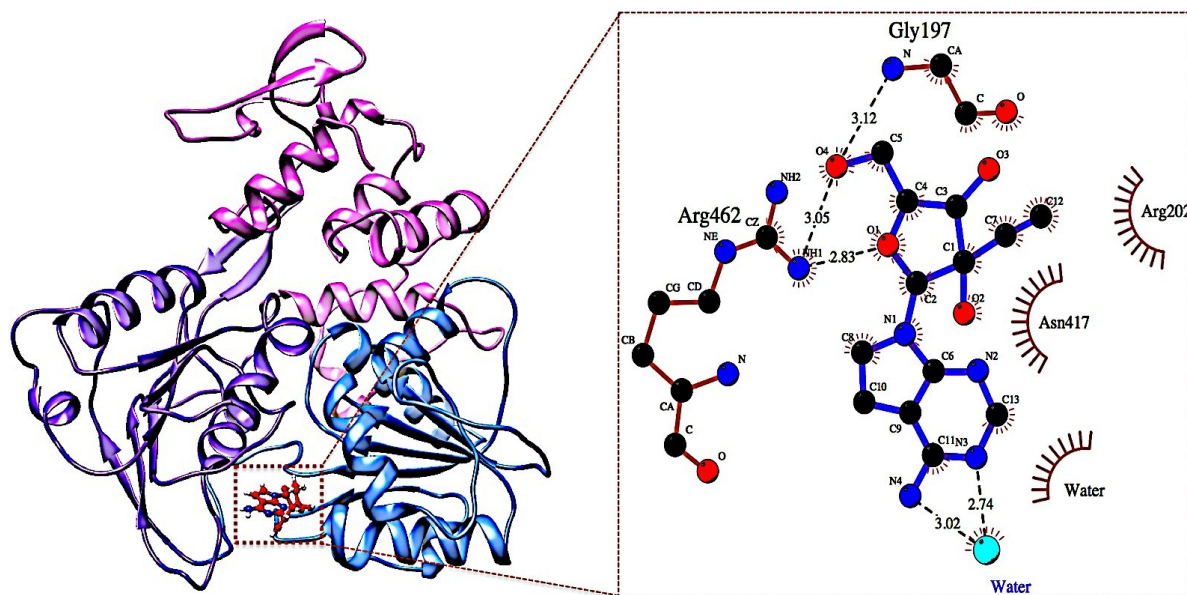


Figure S4: Complex of NITD008-NS3 Helicase with a Docking score of -7.1 kcal/mol. MM/GBSA calculations yielded a result of -11.86 kcal/mol. This ligand showed a similar pose to that of the -7.6 kcal/mol-docked pose, however, there was only one residue, Arg462, which showed stabilizing hydrogen bonds with the terminal oxygen located on the ribose group of NITD008.

Energy Components (kcal/mol)					
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV NS3 Helicase	-3546.66 ± 28.91	-30102.45 ± 107.82	-33649.11 ± 103.65	-4466.57 ± 89.33	-38115.68 ± 48.08
NITD008 (Pose 4)	-5.86 ± 0.44	12.17 ± 4.43	6.31 ± 4.48	-28.40 ± 2.23	-22.09 ± 3.39
Complex	-27.60 ± 3.22	-25.88 ± 6.98	-53.48 ± 6.59	29.49 ± 4.00	-23.99 ± 4.06

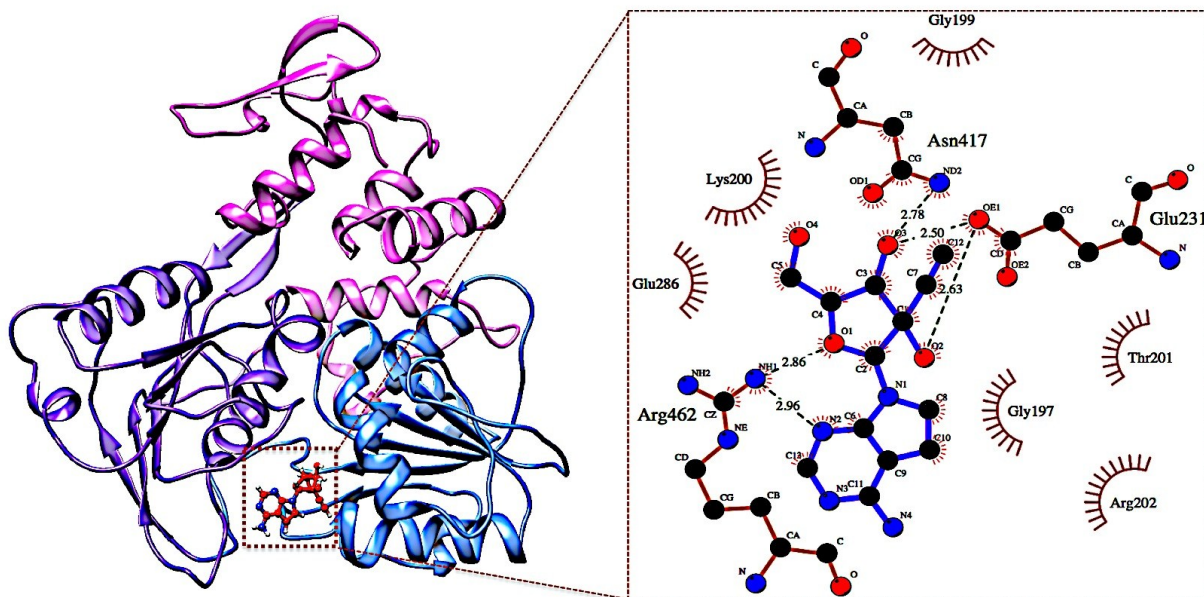


Figure S5: Complex of NITD008- NS3 Helicase with a Docking score of -7.1 kcal/mol. MM/GBSA calculations yielded a result of -23.99 kcal/mol. This pose showed the same docking score as the above ligand, however, three residues: Arg462, Asn417, and Glu231, were involved in stabilizing hydrogen bonds.

Energy Components (kcal/mol)					
	ΔE_{vdW}	ΔE_{elec}	ΔG_{gas}	ΔG_{solv}	ΔG_{bind}
ZIKV NS3 Helicase	-3541.38 ± 29.25	-29907.78 ± 118.22	-33449.15 ± 118.80	-4626.31 ± 114.92	-38075.46 ± 56.93
NITD008 (Pose 5)	-5.54 ± 0.70	12.28 ± 5.43	6.74 ± 5.32	-26.45 ± 2.46	-19.71 ± 4.02
Complex	-13.69 ± 3.83	-7.75 ± 6.43	-21.44 ± 7.73	15.54 ± 5.95	-5.90 ± 3.06

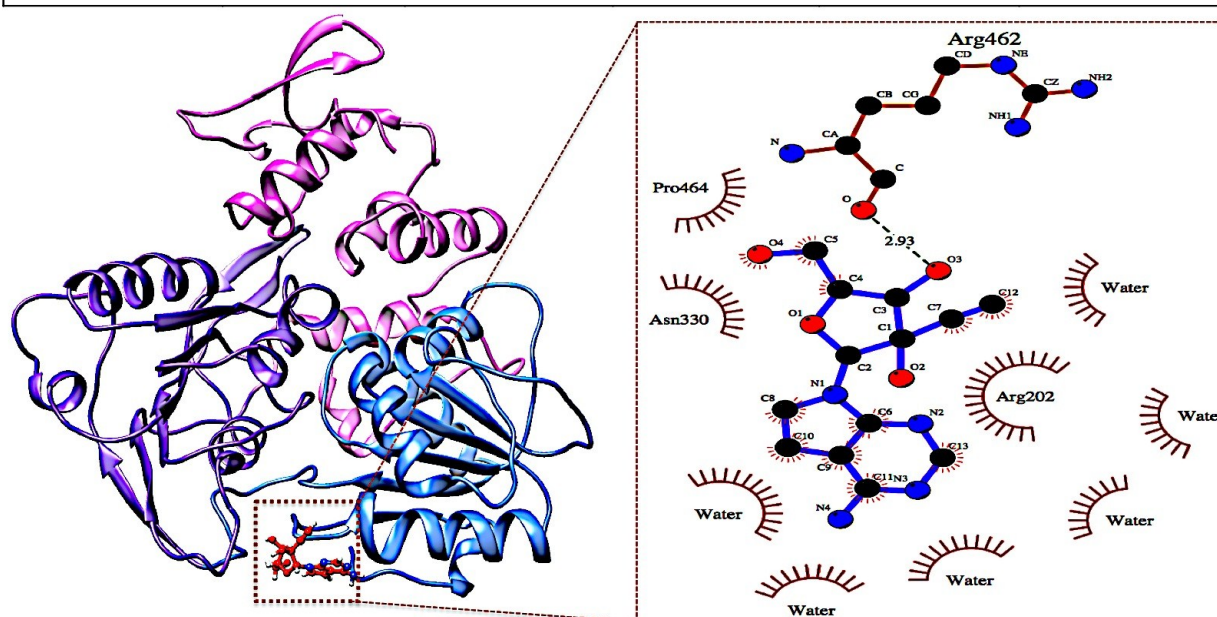


Figure S6: MM/GBSA calculations yielded a result of -5.90 kcal/mol, which was lower than that of the docking score of 6.9 kcal/mol. This was due to the ligand binding out of the active site of the enzyme, thus leading to minimal intermolecular forces at the hydrophobic pocket.